



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

(21) BR 122012021663-1 A2



(22) Data do Depósito: 02/04/2010

(43) Data da Publicação Nacional: 07/10/2010

(54) **Título:** DISPOSITIVO DE CODIFICAÇÃO DE VOZ, DISPOSITIVO DE DECODIFICAÇÃO DE VOZ, MÉTODO DE CODIFICAÇÃO DE VOZ, MÉTODO DE DECODIFICAÇÃO DE VOZ, PROGRAMA DE CODIFICAÇÃO DE VOZ E PROGRAMA DE DECODIFICAÇÃO DE VOZ

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

(30) **Prioridade Unionista:** 03/04/2009 JP 2009-091396; 08/07/2009 JP 2009-162238; 12/01/2010 JP 2010-004419; 19/06/2009 JP 2009-146831.

(71) **Depositante(es):** NTT DOCOMO, INC..

(72) **Inventor(es):** KOSUKE TSUJINO; KEI KIKUIRI; NOBUHIKO NAKA.

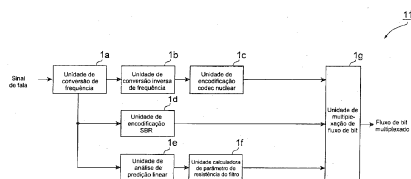
(86) **Pedido PCT:** PCT JP2010056077 de 02/04/2010

(87) **Publicação PCT:** WO 2010/114123 de 07/10/2010

(85) **Data da Fase Nacional:** 28/08/2012

(62) **Pedido original do dividido:** PI1015049-8 - 02/04/2010

(57) **Resumo:** DISPOSITIVO DE CODIFICAÇÃO DE VOZ, DISPOSITIVO DE DECODIFICAÇÃO DE VOZ, MÉTODO DE CODIFICAÇÃO DE VOZ, MÉTODO DE DECODIFICAÇÃO DE VOZ, PROGRAMA DE CODIFICAÇÃO DE VOZ E PROGRAMA DE DECODIFICAÇÃO DE VOZ. Um coeficiente de previsão linear de um sinal representado em um domínio de frequência é representado realizando-se a análise de previsão linear in a direção de frequência com uso de um método de covariância ou um método de auto-correlação. Após a resistência de filtro do coeficiente de previsão linear obtido ser ajustada, a filtragem é realizada na direção de frequência do sinal com uso do coeficiente ajustado, assim o envelope temporal do sinal é transformado. Isto reduz a ocorrência de pré-eco e pós-eco e melhora a qualidade subjetiva do sinal decodificado, sem aumentar significativamente a taxa de bit em uma técnica de extensão de banda no domínio de frequência representado por SBR.



Relatório Descritivo da Patente de Invenção para "**DISPOSITIVO DE CODIFICAÇÃO DE VOZ, DISPOSITIVO DE DECODIFICAÇÃO DE VOZ, MÉTODO DE CODIFICAÇÃO DE VOZ, MÉTODO DE DECODIFICAÇÃO DE VOZ, PROGRAMA DE CODIFICAÇÃO DE VOZ E PROGRAMA**
5 **DE DECODIFICAÇÃO DE VOZ**".

Pedido dividido do PI1015049-8, depositado em 02.04.2010.

Campo técnico

A presente invenção se refere a um dispositivo de codificação de voz, um dispositivo de decodificação de voz, um método de codificação de voz, um método de decodificação de voz, um programa de codificação de voz e um programa de decodificação de voz.
10

Técnica anterior

As técnicas de codificação de áudio de voz para compactar a quantidade de dados de sinais em poucos décimos removendo as informações não necessárias para percepção humana com uso de psicologia auditiva são extremamente importantes na transmissão e armazenamento de sinais. Os exemplos de técnicas de codificação de áudio perceptuais amplamente usadas incluem "MPEG4 AAC" padronizado por "ISO/IEC MPEG".
15

Uma técnica de extensão de banda para gerar componentes de alta frequência com uso de componentes de baixa frequência de voz tem sido amplamente usada recentemente para aumentar o desempenho da codificação de e obter uma alta qualidade em uma baixa taxa de bit. Os exemplos típicos da técnica de extensão de banda incluem a técnica SBR (Replicação de Banda Espectral) usada em "MPEG4 AAC". Na SBR, um componente de alta frequência é gerado convertendo-se um sinal em uma região espectral com uso de um banco de filtro QMF (Filtro em Espelho de Quadratura) e copiando-se os coeficientes espectrais da uma banda de baixa frequência para uma banda de alta frequência com relação ao sinal convertido, e o componente de alta frequência é ajustado ajustando-se o envelope espectral e a tonalidade dos coeficientes copiados. Devido ao fato de que um
20
25
30

método de codificação de voz com uso da técnica de extensão de banda pode reproduzir os componentes de alta frequência de um sinal com uso

somente de uma pequena quantidade de informações suplementares, este é efetivo na redução da taxa de bit da codificação de voz.

Na técnica de extensão de banda no domínio de frequência representado pela SBR, a tonalidade e envelope espectral do coeficiente espectral representado no domínio de frequência são ajustados realizando-se novamente o ajuste, realizando-se a filtragem inversa de previsão linear em uma direção temporal, e sobrepondo-se o ruído no coeficiente espectral. Como resultado deste processo de ajuste, mediante a codificação de um sinal que tem uma grande variação no envelope temporal tal como um sinal de voz, aplausos ou castanholas, um ruído de reverberação chamado de um pré-eco ou um pós-eco pode ser percebido no sinal decodificado. Este problema é causado devido ao fato de que o envelope temporal do componente de alta frequência é transformado durante o processo de ajuste, e em muitos casos, o envelope temporal é mais suave depois do processo de ajuste do que antes do processo de ajuste. O envelope temporal do componente de alta frequência após o processo de ajuste não combina com o envelope temporal do componente de alta frequência de um sinal original antes de ser codificado, assim causando o pré-eco e o pós-eco.

Um problema similar àquele do pré-eco e pós-eco também ocorre na codificação de áudio de múltiplos canais com uso de um processo paramétrico representado pelo "MPEG Surround" e um estéreo paramétrico. Um decodificador usado na codificação de áudio de múltiplos canais inclui um meio para realizar a decorrelação em um sinal decodificado com uso de um filtro de reverberação. No entanto, o envelope temporal do sinal é transformado durante a decorrelação, assim causando degradação de um sinal de reprodução similar àquele do pré-eco e pós-eco. As soluções para o problema incluem uma técnica de TES (Formatação de Envelope Temporal) (Literatura de Patente 1). Na técnica de TES, uma análise de previsão linear é realizada em uma direção de frequência em um sinal representado em um domínio de QMF em que a decorrelação não foi ainda realizada para se obter um coeficiente de previsão linear, e, com uso do coeficiente de previsão linear, a filtragem de síntese de previsão linear é realizada na direção de fre-

quência no sinal em que a decorrelação foi realizada. Este processo permite que a técnica de TES extraia o envelope temporal de um sinal em que a decorrelação ainda não foi realizada, e em concordância com o envelope temporal extraído, ajustar o envelope temporal do sinal em que a decorrelação foi realizada. Devido ao fato de que o sinal em que a decorrelação não foi ainda realizada tem um envelope temporal menos distorcido, o envelope temporal do sinal em que a decorrelação foi realizada é ajustado para um formato menos distorcido, assim obtendo-se um sinal de reprodução em que o pré-eco e o pós-eco são melhorados.

10 **Lista de Citação**

Literatura de Patente

Literatura de Patente 1: Publicação de Pedido de Patente dos Estados Unidos N° 2006/0239473

Sumário da Invenção

15 Problema Técnico

A técnica de TES descrita acima é uma técnica que utiliza o fato de que um sinal em que a decorrelação não foi realizada ainda tem um envelope temporal menos distorcido. No entanto, em um decodificador de SBR, o componente de alta frequência de um sinal é copiado do componente de baixa frequência do sinal. Da mesma forma, não é possível se obter um envelope temporal menos distorcido com relação ao componente de alta frequência. Uma das soluções para este problema é um método para analisar o componente de alta frequência de um sinal de entrada em um codificado de SBR, quantizar o coeficiente de previsão linear obtido como resultado da análise, e multiplexá-los em um fluxo de bit a ser transmitido. Este método permite que o decodificador de SBR possa obter um coeficiente de previsão linear que inclui informações com envelope temporal menos distorcido do componente de alta frequência. No entanto, neste caso, uma grande quantidade de informações é necessária para transmitir o coeficiente de previsão linear quantizado, assim aumentando significativamente a taxa de bit de todo o fluxo de bit codificado. Assim, a presente invenção pretende reduzir a ocorrência de pré-eco e pós-eco e melhorar a qualidade subjetiva do

5 sinal decodificado, sem aumentar significativamente a taxa de bit na técnica de extensão de banda no domínio de frequência representado pela SBR.

Solução do Problema

5 Um dispositivo de codificação de voz da presente invenção é um dispositivo de codificação de voz para codificar um sinal de voz e inclui: meio de codificação central para codificar um componente de baixa frequência do sinal de voz; meio de cálculo de informações suplementares de envelope temporal para calcular as informações suplementares de envelope temporal para se obter uma aproximação de um envelope temporal de um componen-

10 te de alta frequência do sinal de voz com uso de um envelope temporal do componente de baixa frequência do sinal de voz; e meio de multiplexação de fluxo de bit para gerar um fluxo de bit em que ao menos o componente de baixa frequência codificado pelo meio de codificação central e as informa-

15 ções suplementares de envelope temporal calculadas pelo meio de cálculo de informações suplementares de envelope temporal são multiplexados.

20 No dispositivo de codificação de voz da presente invenção, as informações suplementares de envelope temporal preferencialmente representam um parâmetro que indica uma sutileza de variação no envelope temporal do componente de alta frequência do sinal de voz em uma seção de análise predeterminada.

25 É preferível que o dispositivo de codificação de voz da presente invenção ainda inclua um meio de conversão de frequência para converter o sinal de voz em um domínio de frequência, e o meio de cálculo de informações suplementares de envelope temporal calcule as informações suplementares de envelope temporal com base em um coeficiente de previsão linear de alta frequência obtido realizando-se a análise de previsão linear em uma direção de frequência em um coeficiente de lado de frequência do sinal de voz convertido no domínio de frequência pelo meio de conversão de frequência.

30 No dispositivo de codificação de voz da presente invenção, o meio de cálculo de informações suplementares de envelope temporal preferencialmente realiza a análise de previsão linear em uma direção de fre-

quência em um coeficiente de lado de baixa frequência do sinal de voz convertido no domínio de frequência pelo meio de conversão de frequência para se obter um coeficiente de previsão linear de baixa frequência, e calcula as informações suplementares de envelope temporal com base no coeficiente
5 de previsão linear de baixa frequência e no coeficiente de previsão linear de alta frequência.

No dispositivo de codificação de voz da presente invenção, o meio de cálculo de informações suplementares de envelope temporal preferencialmente obtém um ganho de previsão a partir de cada um dos coeficientes de previsão linear de baixa frequência e dos coeficientes de previsão
10 linear de alta frequência, e calcula as informações suplementares de envelope temporal com base nas magnitudes dos dois ganhos de previsão.

No dispositivo de codificação de voz da presente invenção, o meio de cálculo de informações suplementares de envelope temporal preferencialmente separa o componente de alta frequência do sinal de voz, obtém
15 as informações de envelope temporal representadas em um domínio de tempo do componente de alta frequência, e calcula as informações suplementares de envelope temporal com base na magnitude da variação temporal das informações de envelope temporal.

No dispositivo de codificação de voz da presente invenção, as informações suplementares de envelope temporal preferencialmente incluem informações diferenciais para obter um coeficiente de previsão linear de alta frequência com uso de um coeficiente de previsão linear de baixa frequência obtido realizando-se a análise de previsão linear em uma direção de frequência no componente de baixa frequência do sinal de voz.
20
25

É preferível que o dispositivo de codificação de voz da presente invenção ainda inclua um meio de conversão de frequência para converter o sinal de voz em um domínio de frequência, e o meio de cálculo de informações suplementares de envelope temporal realize análise de previsão linear
30 em uma direção de frequência em cada um dentre o componente de baixa frequência e o componente de alta frequência do sinal de voz convertido no domínio de frequência pelo meio de conversão de frequência para obter um

coeficiente de previsão linear de baixa frequência e um coeficiente de previsão linear de alta frequência, e obter as informações diferenciais obtendo-se uma diferença entre o coeficiente de previsão linear de baixa frequência e o coeficiente de previsão linear de alta frequência.

- 5 No dispositivo de codificação de voz da presente invenção, as informações diferenciais preferencialmente representam uma diferença entre coeficientes de previsão linear em ao menos qualquer domínio dentre LSP (Par de Espectros Lineares), ISP (Par de Espectros de Imitância), LSF (Frequência de Espectro Linear), ISF (Frequência de Espectro de Imitância) e
- 10 coeficiente PARCOR.

- Um dispositivo de codificação de voz da presente invenção é um dispositivo de codificação de voz para codificar um sinal de voz e inclui: meio de codificação central para codificar um componente de baixa frequência do sinal de voz; meio de conversão de frequência para converter o sinal de voz
- 15 em um domínio de frequência; meio de análise de previsão linear para realizar análise de previsão linear em uma direção de frequência em um coeficiente de lado de alta frequência do sinal de voz convertido no domínio de frequência pelo meio de conversão de frequência para obter um coeficiente de previsão linear de alta frequência; meio de decimação de coeficiente de pre-
- 20 visão para decimar o coeficiente de previsão linear de alta frequência obtido pelo meio análise de previsão linear em uma direção temporal; meio de quantização de coeficiente de previsão para quantizar o coeficiente de previsão linear de alta frequência decimado pelo meio de decimação de coeficiente de previsão; e meio de multiplexação de fluxo de bit para gerar um fluxo
- 25 de bit em que ao menos o componente de baixa frequência codificado pelo meio de codificação central e o coeficiente de previsão linear de alta frequência quantizado pelo meio de quantização de coeficiente de previsão são multiplexados.

- Um dispositivo de decodificação de voz da presente invenção é
- 30 um dispositivo de decodificação de voz para decodificar um sinal de voz codificado e inclui: meio de separação de fluxo de bit para separar um fluxo de bit recebido de fora do dispositivo de decodificação de voz que inclui o sinal

de voz codificado em um fluxo de bit codificado e informações suplementares de envelope temporal; meio de decodificação central para decodificar o fluxo de bit codificado separado pelo meio de separação de fluxo de bit para obter um componente de baixa frequência; meio de conversão de frequência para converter o componente de baixa frequência obtido pelo meio de decodificação central para um domínio de frequência; meio de geração de alta frequência para gerar um componente de alta frequência copiando o componente de baixa frequência convertido no domínio de frequência pelo meio de conversão de frequência de uma banda de baixa frequência em uma banda de alta frequência; meio de análise de envelope temporal de baixa frequência para analisar o componente de baixa frequência convertido no domínio de frequência pelo meio de conversão de frequência para obter as informações de envelope temporal; meio de ajuste de envelope temporal para ajustar as informações de envelope temporal obtidas pelo meio de análise de envelope temporal de baixa frequência com uso das informações suplementares de envelope temporal, e meio de transformação de envelope temporal para transformar um envelope temporal do componente de alta frequência gerado pelo meio de geração de alta frequência com uso das informações de envelope temporal ajustadas pelo meio de ajuste de envelope temporal.

É preferível que o dispositivo de decodificação de voz da presente invenção ainda inclua meio de ajuste de alta frequência para ajustar o componente de alta frequência, e o meio de conversão de frequência pode ser um banco de filtro QMF de 64 divisões com um coeficiente real ou complexo, e o meio de conversão de frequência, o meio de geração de alta frequência e o meio de ajuste de alta frequência operam com base em um decodificador de Replicação de Banda Espectral (SBR) para o "MPEG4 AAC" definido em "ISO/IEC 14496-3".

No dispositivo de decodificação de voz da presente invenção, é preferível que o meio de análise de envelope temporal de baixa frequência realize a análise de previsão linear em uma direção de frequência no componente de baixa frequência convertido no domínio de frequência pelo meio de conversão de frequência para obter um coeficiente de previsão linear de

baixa frequência, o meio de ajuste de envelope temporal pode ajustar o coeficiente de previsão linear de baixa frequência com uso das informações suplementares de envelope temporal, e o meio de transformação de envelope temporal pode realizar filtragem de previsão linear em uma direção de frequência no componente de alta frequência no domínio de frequência gerado pelo meio de geração de alta frequência, com uso de um coeficiente de previsão linear ajustado pelo meio de ajuste de envelope temporal, para transformar um envelope temporal de um sinal de voz.

No dispositivo de decodificação de voz da presente invenção, é preferível que o meio de análise de envelope temporal de baixa frequência obtenha as informações de envelope temporal de um sinal de voz obtendo força de cada abertura de tempo do componente de baixa frequência convertido no domínio de frequência pelo meio de conversão de frequência, o meio de ajuste de envelope temporal ajuste as informações de envelope temporal com uso das informações suplementares de envelope temporal, e o meio de transformação de envelope temporal sobreponha as informações de envelope temporal ajustadas no componente de alta frequência no domínio de frequência gerado pelo meio de geração de alta frequência para transformar um envelope temporal de um componente de alta frequência.

No dispositivo de decodificação de voz da presente invenção, é preferível que o meio de análise de envelope temporal de baixa frequência obtenha informações de envelope temporal de um sinal de voz obtendo força de cada amostra de sub-banda QMF do componente de baixa frequência convertido no domínio de frequência pelo meio de conversão de frequência, o meio de ajuste de envelope temporal ajuste as informações de envelope temporal com uso das informações suplementares de envelope temporal, e o meio de transformação de envelope temporal transforme um envelope temporal de um componente de alta frequência multiplicando as informações de envelope temporal ajustadas pelo componente de alta frequência no domínio de frequência gerado pelo meio de geração de alta frequência.

No dispositivo de decodificação de voz da presente invenção, as informações suplementares de envelope temporal preferencialmente repre-

sentam um parâmetro de resistência de filtro a usado para ajustar a resistência de um coeficiente de previsão linear.

No dispositivo de decodificação de voz da presente invenção, as informações suplementares de envelope temporal preferencialmente representam um parâmetro que indica a magnitude da variação temporal das informações de envelope temporal.

No dispositivo de decodificação de voz da presente invenção, as informações suplementares de envelope temporal preferencialmente incluem informações diferenciais de um coeficiente de previsão linear com relação ao coeficiente de previsão linear de baixa frequência.

No dispositivo de decodificação de voz da presente invenção, as informações diferenciais preferencialmente representam uma diferença entre os coeficientes de previsão linear em ao menos qualquer domínio dentre LSP (Par de Espectros Lineares), ISP (Par de Espectros de Imitância), LSF (Frequência de Espectro Linear), ISF (Frequência de Espectro de Imitância), e coeficiente PARCOR.

No dispositivo de decodificação de voz da presente invenção, é preferível que o meio de análise de envelope temporal de baixa frequência realize a análise de previsão linear em uma direção de frequência no componente de baixa frequência convertido no domínio de frequência pelo meio de conversão de frequência para obter o coeficiente de previsão linear de baixa frequência, e obter força de cada abertura de tempo do componente de baixa frequência no domínio de frequência para obter as informações de envelope temporal de um sinal de voz, o meio de ajuste de envelope temporal ajuste o coeficiente de previsão linear de baixa frequência com uso das informações suplementares de envelope temporal e ajuste as informações de envelope temporal com uso das informações suplementares de envelope temporal, e o meio de transformação de envelope temporal realize a filtragem de previsão linear em uma direção de frequência no componente de alta frequência no domínio de frequência gerado pelo meio de geração de alta frequência com uso do coeficiente de previsão linear ajustado pelo meio de ajuste de envelope temporal para transformar um envelope temporal de

um sinal de voz, e transforme um envelope temporal do componente de alta frequência sobrepondo as informações de envelope temporal ajustadas pelo meio de ajuste de envelope temporal no componente de alta frequência no domínio de frequência.

5 No dispositivo de decodificação de voz da presente invenção, é preferível que o meio de análise de envelope temporal de baixa frequência realize a análise de previsão linear em uma direção de frequência no componente de baixa frequência convertido no domínio de frequência pelo meio de conversão de frequência para obter o coeficiente de previsão linear de
10 baixa frequência, e obtenha informações de envelope temporal de um sinal de voz obtendo força de cada amostra de sub-banda QMF do componente de baixa frequência no domínio de frequência, o meio de ajuste de envelope temporal ajuste o coeficiente de previsão linear de baixa frequência com uso das informações suplementares de envelope temporal e ajuste as informa-
15 ções de envelope temporal com uso das informações suplementares de envelope temporal, e o meio de transformação de envelope temporal realize filtragem de previsão linear em uma direção de frequência em um componente de alta frequência no domínio de frequência gerado pelo meio de geração de alta frequência com uso de um coeficiente de previsão linear ajustado pelo meio de ajuste de envelope temporal para transformar um envelo-
20 pe temporal de um sinal de voz, e transforme um envelope temporal do componente de alta frequência multiplicando as informações de envelope temporal ajustadas pelo meio de ajuste de envelope temporal pelo componente de alta frequência no domínio de frequência.

25 No dispositivo de decodificação de voz da presente invenção, as informações suplementares de envelope temporal preferencialmente representam um parâmetro que indica tanto uma resistência de filtro de um coeficiente de previsão linear como uma magnitude de variação temporal das informações de envelope temporal.

30 Um dispositivo de decodificação de voz da presente invenção é um dispositivo de decodificação de voz para decodificar um sinal de voz codificado e inclui: meio de separação de fluxo de bit para separar um fluxo de

bit recebido de fora do dispositivo de decodificação de voz que inclui o sinal de voz codificado em um fluxo de bit codificado e um coeficiente de previsão linear, coeficiente de previsão linear meio de interpolação/extrapolação para interpolar ou extrapolar o coeficiente de previsão linear em uma direção temporal, e meio de transformação de envelope temporal para realizar a filtragem de previsão linear em uma direção de frequência em um componente de alta frequência representado em um domínio de frequência com uso de um coeficiente de previsão linear interpolado ou extrapolado pelo meio de interpolação/extrapolação de coeficiente de previsão linear para transformar um envelope temporal de um sinal de voz.

Um método de codificação de voz da presente invenção é um método de codificação de voz que usa um dispositivo de codificação de voz para codificar um sinal de voz e inclui: uma etapa de codificação central em que o dispositivo de codificação de voz codifica um componente de baixa frequência do sinal de voz; uma etapa de conversão de frequência em que o dispositivo de codificação de voz converte o sinal de voz em um domínio de frequência; uma etapa de análise de previsão linear em que o dispositivo de codificação de voz obtém coeficientes de previsão linear de alta frequência realizando a análise de previsão linear em uma direção de frequência em coeficientes de lado de alta frequência do sinal de voz convertido no domínio de frequência na etapa de conversão de frequência; uma etapa de decimação de coeficiente de previsão em que o dispositivo de codificação de voz decima o coeficiente de previsão linear de alta frequência obtido na etapa de análise de previsão linear em uma direção temporal; uma etapa de quantização de coeficiente de previsão em que o dispositivo de codificação de voz quantiza os coeficientes de previsão linear de alta frequência decimado na etapa de decimação de coeficiente de previsão; e uma etapa de multiplexação de fluxo de bit em que o dispositivo de codificação de voz gera um fluxo de bit em que ao menos o componente de baixa frequência codificado na etapa de codificação central e os coeficientes de previsão linear de alta frequência quantizados na etapa de quantização de coeficiente de previsão são multiplexados.

Um dispositivo de decodificação de voz da presente invenção é um dispositivo de decodificação de voz para decodificar um sinal de voz codificado e inclui: meio de separação de fluxo de bit para separar um fluxo de bit recebido de fora do dispositivo de decodificação de voz que inclui o sinal
5 de voz codificado em um fluxo de bit codificado e informações suplementares de envelope temporal; meio de decodificação central para decodificar o fluxo de bit codificado separado pelo meio de separação de fluxo de bit para obter um componente de baixa frequência; meio de conversão de frequência para converter o componente de baixa frequência obtido pelo meio de decodificação central para um domínio de frequência; meio de geração de alta
10 frequência para gerar um componente de alta frequência copiando o componente de baixa frequência convertido no domínio de frequência pelo meio de conversão de frequência de uma banda de baixa frequência em uma banda de alta frequência; meio de análise de envelope temporal de baixa frequência para analisar o componente de baixa frequência convertido no domínio
15 de frequência pelo meio de conversão de frequência para obter as informações de envelope temporal; meio de ajuste de envelope temporal para ajustar as informações de envelope temporal obtidas pelo meio de análise de envelope temporal de baixa frequência com uso das informações suplementares de envelope temporal, e meio de transformação de envelope temporal
20 para transformar um envelope temporal do componente de alta frequência gerado pelo meio de geração de alta frequência com uso das informações de envelope temporal ajustadas pelo meio de ajuste de envelope temporal.

É preferível que o dispositivo de decodificação de voz da presente invenção ainda inclua meio de ajuste de alta frequência para ajustar o
25 componente de alta frequência, e o meio de conversão de frequência pode ser um banco de filtro QMF de 64 divisões com um coeficiente real ou complexo, e o meio de conversão de frequência, o meio de geração de alta frequência e o meio de ajuste de alta frequência operam com base em um decodificador de Replicação de Banda Espectral (SBR) para o "MPEG4 AAC"
30 definido em "ISO/IEC 14496-3".

No dispositivo de decodificação de voz da presente invenção, é preferível que o meio de análise de envelope temporal de baixa frequência realize a análise de previsão linear em uma direção de frequência no componente de baixa frequência convertido no domínio de frequência pelo meio de conversão de frequência para obter um coeficiente de previsão linear de baixa frequência, o meio de ajuste de envelope temporal pode ajustar o coeficiente de previsão linear de baixa frequência com uso das informações suplementares de envelope temporal, e o meio de transformação de envelope temporal pode realizar filtragem de previsão linear em uma direção de frequência no componente de alta frequência no domínio de frequência gerado pelo meio de geração de alta frequência, com uso de um coeficiente de previsão linear ajustado pelo meio de ajuste de envelope temporal, para transformar um envelope temporal de um sinal de voz.

No dispositivo de decodificação de voz da presente invenção, é preferível que o meio de análise de envelope temporal de baixa frequência obtenha as informações de envelope temporal de um sinal de voz obtendo força de cada abertura de tempo do componente de baixa frequência convertido no domínio de frequência pelo meio de conversão de frequência, o meio de ajuste de envelope temporal ajuste as informações de envelope temporal com uso das informações suplementares de envelope temporal, e o meio de transformação de envelope temporal sobreponha as informações de envelope temporal ajustadas no componente de alta frequência no domínio de frequência gerado pelo meio de geração de alta frequência para transformar um envelope temporal de um componente de alta frequência.

No dispositivo de decodificação de voz da presente invenção, é preferível que o meio de análise de envelope temporal de baixa frequência obtenha informações de envelope temporal de um sinal de voz obtendo força de cada amostra de sub-banda QMF do componente de baixa frequência convertido no domínio de frequência pelo meio de conversão de frequência, o meio de ajuste de envelope temporal ajuste as informações de envelope temporal com uso das informações suplementares de envelope temporal, e o meio de transformação de envelope temporal transforme um envelope tem-

poral de um componente de alta frequência multiplicando as informações de envelope temporal ajustadas pelo componente de alta frequência no domínio de frequência gerado pelo meio de geração de alta frequência.

No dispositivo de decodificação de voz da presente invenção, as
5 informações suplementares de envelope temporal preferencialmente representam um parâmetro de resistência de filtro a usado para ajustar a resistência de um coeficiente de previsão linear.

No dispositivo de decodificação de voz da presente invenção, as
informações suplementares de envelope temporal preferencialmente repre-
10 sentam um parâmetro que indica a magnitude da variação temporal das informações de envelope temporal.

No dispositivo de decodificação de voz da presente invenção, as
informações suplementares de envelope temporal preferencialmente inclu-
des informações diferenciais de um coeficiente de previsão linear com rela-
15 ção ao coeficiente de previsão linear de baixa frequência.

No dispositivo de decodificação de voz da presente invenção, as
informações diferenciais preferencialmente representam uma diferença entre
os coeficientes de previsão linear em ao menos qualquer domínio dentre
LSP (Par de Espectros Lineares), ISP (Par de Espectros de Imitância), LSF
20 (Frequência de Espectro Linear), ISF (Frequência de Espectro de Imitância),
e coeficiente PARCOR.

No dispositivo de decodificação de voz da presente invenção, é
preferível que o meio de análise de envelope temporal de baixa frequência
realize a análise de previsão linear em uma direção de frequência no com-
25 ponente de baixa frequência convertido no domínio de frequência pelo meio
de conversão de frequência para obter o coeficiente de previsão linear de
baixa frequência, e obter força de cada abertura de tempo do componente
de baixa frequência no domínio de frequência para obter as informações de
envelope temporal de um sinal de voz, o meio de ajuste de envelope tempo-
30 ral ajuste o coeficiente de previsão linear de baixa frequência com uso das
informações suplementares de envelope temporal e ajuste as informações
de envelope temporal com uso das informações suplementares de envelope

- temporal, e o meio de transformação de envelope temporal realize a filtragem de previsão linear em uma direção de frequência no componente de alta frequência no domínio de frequência gerado pelo meio de geração de alta frequência com uso do coeficiente de previsão linear ajustado pelo meio
- 5 de ajuste de envelope temporal para transformar um envelope temporal de um sinal de voz, e transforme um envelope temporal do componente de alta frequência sobrepondo as informações de envelope temporal ajustadas pelo meio de ajuste de envelope temporal no componente de alta frequência no domínio de frequência.
- 10 No dispositivo de decodificação de voz da presente invenção, é preferível que o meio de análise de envelope temporal de baixa frequência realize a análise de previsão linear em uma direção de frequência no componente de baixa frequência convertido no domínio de frequência pelo meio de conversão de frequência para obter o coeficiente de previsão linear de
- 15 baixa frequência, e obtenha informações de envelope temporal de um sinal de voz obtendo força de cada amostra de sub-banda QMF do componente de baixa frequência no domínio de frequência, o meio de ajuste de envelope temporal ajuste o coeficiente de previsão linear de baixa frequência com uso das informações suplementares de envelope temporal e ajuste as informa-
- 20 ções de envelope temporal com uso das informações suplementares de envelope temporal, e o meio de transformação de envelope temporal realize filtragem de previsão linear em uma direção de frequência em um componente de alta frequência no domínio de frequência gerado pelo meio de geração de alta frequência com uso de um coeficiente de previsão linear ajustado pelo meio de ajuste de envelope temporal para transformar um envelo-
- 25 pe temporal de um sinal de voz, e transforme um envelope temporal do componente de alta frequência multiplicando as informações de envelope temporal ajustadas pelo meio de ajuste de envelope temporal pelo componente de alta frequência no domínio de frequência.
- 30 No dispositivo de decodificação de voz da presente invenção, as informações suplementares de envelope temporal preferencialmente representam um parâmetro que indica tanto uma resistência de filtro de um coefi-

ciente de previsão linear como uma magnitude de variação temporal das informações de envelope temporal.

Um dispositivo de decodificação de voz da presente invenção é um dispositivo de decodificação de voz para decodificar um sinal de voz codificado e inclui: meio de separação de fluxo de bit para separar um fluxo de bit recebido de fora do dispositivo de decodificação de voz que inclui o sinal de voz codificado em um fluxo de bit codificado e um coeficiente de previsão linear, coeficiente de previsão linear meio de interpolação/extrapolação para interpolar ou extrapolar o coeficiente de previsão linear em uma direção temporal, e meio de transformação de envelope temporal para realizar a filtragem de previsão linear em uma direção de frequência em um componente de alta frequência representado em um domínio de frequência com uso de um coeficiente de previsão linear interpolado ou extrapolado pelo meio de interpolação/extrapolação de coeficiente de previsão linear para transformar um envelope temporal de um sinal de voz.

Um método de codificação de voz da presente invenção é um método de codificação de voz que usa um dispositivo de codificação de voz para codificar um sinal de voz e inclui: uma etapa de codificação núcleo em que o dispositivo de codificação de voz codifica um componente de baixa frequência do sinal de voz; uma etapa de cálculo de informações suplementares de envelope temporal em que o dispositivo de codificação de voz calcula informações suplementares de envelope temporal para obter uma aproximação de um envelope temporal de um componente de alta frequência do sinal de voz com uso de um envelope temporal de um componente de baixa frequência do sinal de voz; e uma etapa de multiplexação de fluxo de bit em que o dispositivo de codificação de voz gera um fluxo de bit em que ao menos o componente de baixa frequência codificado na etapa de codificação central e as informações suplementares de envelope temporal calculadas na etapa de cálculo de informações suplementares de envelope temporal são multiplexados.

Um método de codificação de voz da presente invenção é um método de codificação de voz que usa um dispositivo de codificação de voz

para codificar um sinal de voz e inclui: uma etapa de codificação central em que o dispositivo de codificação de voz codifica um componente de baixa frequência do sinal de voz; uma etapa de conversão de frequência em que o dispositivo de codificação de voz converte o sinal de voz em um domínio de frequência; uma etapa de análise de previsão linear em que o dispositivo de codificação de voz obtém coeficientes de previsão linear de alta frequência realizando a análise de previsão linear em uma direção de frequência em coeficientes de lado de alta frequência do sinal de voz convertido no domínio de frequência na etapa de conversão de frequência; uma etapa de decimação de coeficiente de previsão em que o dispositivo de codificação de voz decima o coeficiente de previsão linear de alta frequência obtido na etapa de meio de análise de previsão linear em uma direção temporal; uma etapa de quantização de coeficiente de previsão em que o dispositivo de codificação de voz quantiza os coeficientes de previsão linear de alta frequência decimados na etapa de meio de decimação de coeficiente de previsão; e uma etapa de multiplexação de fluxo de bit em que o dispositivo de codificação de voz gera um fluxo de bit em que ao menos o componente de baixa frequência codificado na etapa de codificação central e os coeficientes de previsão linear de alta frequência quantizados na etapa de quantização de coeficiente de previsão são multiplexados.

Um método de decodificação de voz da presente invenção é um método de decodificação de voz que usa um dispositivo de decodificação de voz para decodificar um sinal de voz codificado e inclui: uma etapa de separação de fluxo de bit em que o dispositivo de decodificação de voz separa um fluxo de bit recebido de fora do dispositivo de decodificação de voz que inclui o sinal de voz codificado em um fluxo de bit codificado e informações suplementares de envelope temporal; uma etapa de decodificação central em que o dispositivo de decodificação de voz obtém um componente de baixa frequência decodificando o fluxo de bit codificado separado na etapa de separação de fluxo de bit; uma etapa de conversão de frequência em que o dispositivo de decodificação de voz converte o componente de baixa frequência obtido na etapa de decodificação central em um domínio de frequência; uma

etapa de geração de alta frequência em que o dispositivo de decodificação de voz gera um componente de alta frequência copiando o componente de baixa frequência convertido no domínio de frequência na etapa de conversão de frequência de uma banda de baixa frequência para uma banda de alta frequência; uma etapa de análise de envelope temporal de baixa frequência em que o dispositivo de decodificação de voz obtém informações de envelope temporal analisando o componente de baixa frequência convertido no domínio de frequência na etapa de conversão de frequência; uma etapa de ajuste de envelope temporal em que o dispositivo de decodificação de voz ajusta as informações de envelope temporal obtidas na etapa de análise de envelope temporal de baixa frequência com uso das informações suplementares de envelope temporal; e uma etapa de transformação de envelope temporal em que o dispositivo de decodificação de voz transforma um envelope temporal do componente de alta frequência gerado na etapa de geração de alta frequência com uso das informações de envelope temporal ajustadas na etapa de ajuste de envelope temporal.

Um método de decodificação de voz da presente invenção é um método de decodificação de voz que usa um dispositivo de decodificação de voz for que decodifica um sinal de voz codificado e inclui: uma etapa de separação de fluxo de bit em que o dispositivo de decodificação de voz separa um fluxo de bit recebido de fora do dispositivo de decodificação de voz que inclui o sinal de voz codificado em um fluxo de bit codificado e um coeficiente de previsão linear; uma etapa de interpolação/extrapolação de coeficiente de previsão linear em que o dispositivo de decodificação de voz interpola ou extrapola o coeficiente de previsão linear em uma direção temporal; e uma etapa de transformação de envelope temporal em que o dispositivo de decodificação de voz transforma um envelope temporal de um sinal de voz realizando a filtragem de previsão linear em uma direção de frequência em um componente de alta frequência representado em um domínio de frequência com uso do coeficiente de previsão linear interpolado ou extrapolado na etapa de interpolação/extrapolação de coeficiente de previsão linear.

Um programa de codificação de voz da presente invenção para codificar um sinal de voz faz com que um dispositivo de computador funcione como: meio de codificação central para codificar um componente de baixa frequência do sinal de voz; meio de cálculo de informações suplementares de envelope temporal para calcular informações suplementares de envelope temporal para obter uma aproximação de um envelope temporal de um componente de alta frequência do sinal de voz com uso de um envelope temporal do componente de baixa frequência do sinal de voz; e meio de multiplexação de fluxo de bit para gerar um fluxo de bit em que ao menos o

5

10

componente de baixa frequência codificado pelo meio de codificação central e as informações suplementares de envelope temporal calculadas pelo meio de cálculo de informações suplementares de envelope temporal são multiplexados.

Um programa de codificação de voz da presente invenção para codificar um sinal de voz que faz com que um dispositivo de computador funcione como: meio de codificação central para codificar um componente de baixa frequência do sinal de voz; meio de conversão de frequência para converter o sinal de voz em um domínio de frequência; meio de análise de previsão linear para realizar a análise de previsão linear em uma direção de

15

20

25

30

frequência em coeficientes de lado de alta frequência do sinal de voz convertido no domínio de frequência pelo meio de conversão de frequência para obter coeficientes de previsão linear de alta frequência; meio de decimação de coeficiente de previsão para decimar coeficientes de previsão linear de alta frequência obtidos pelo meio de análise de previsão linear em uma direção temporal; meio de quantização de coeficiente de previsão para quantizar o coeficiente de previsão linear de alta frequências decimado pelo meio de decimação de coeficiente de previsão; e meio de multiplexação de fluxo de bit para gerar um fluxo de bit em que ao menos o componente de baixa frequência codificado pelo meio de codificação central e coeficientes de previsão linear de alta frequência quantizados pelo meio de quantização de coeficiente de previsão são multiplexados.

Um programa de decodificação de voz da presente invenção para decodificar um sinal de voz codificado faz com que um dispositivo de computador funcione como: meio de separação de fluxo de bit para separar um fluxo de bit recebido de fora do programa de decodificação de voz que inclui o sinal de voz codificado em um fluxo de bit codificado e informações suplementares de envelope temporal; meio de decodificação central para decodificar o fluxo de bit codificado separado pelo meio de separação de fluxo de bit para obter um componente de baixa frequência; meio de conversão de frequência para converter o componente de baixa frequência obtido pelo meio de decodificação central em um domínio de frequência; meio de geração de alta frequência para gerar um componente de alta frequência copiando o componente de baixa frequência convertido no domínio de frequência pelo meio de conversão de frequência de uma banda de baixa frequência para uma banda de alta frequência; meio de análise de envelope temporal de baixa frequência para analisar o componente de baixa frequência convertido no domínio de frequência pelo meio de conversão de frequência para obter informações de envelope temporal; meio de ajuste de envelope temporal para ajustar as informações de envelope temporal obtidas pelo meio de análise de envelope temporal de baixa frequência com uso das informações suplementares de envelope temporal; e meio de transformação de envelope temporal para transformar um envelope temporal do componente de alta frequência gerado pelo meio de geração de alta frequência com uso das informações de envelope temporal ajustadas pelo meio de ajuste de envelope temporal.

Um programa de decodificação de voz da presente invenção para decodificar um sinal de voz codificado faz com que um dispositivo de computador funcione como: meio de separação de fluxo de bit para separar um fluxo de bit que inclui o sinal de voz codificado em um fluxo de bit codificado e um coeficiente de previsão linear. O fluxo de bit recebido de fora do programa de decodificação de voz. Em adição, o programa de decodificação de voz ainda faz com que um dispositivo de computador funcione como; meio de interpolação/extrapolação de coeficiente de previsão linear para in-

terpolar ou extrapolar o coeficiente de previsão linear em uma direção temporal; e meio de transformação de envelope temporal para realizar a filtragem de previsão linear em uma direção de frequência em um componente de alta frequência representado em um domínio de frequência com uso de um coeficiente de previsão linear interpolado ou extrapolado pelo meio de interpolação/extrapolação de coeficiente de previsão linear para transformar um envelope temporal de um sinal de voz.

No dispositivo de decodificação de voz da presente invenção, o meio de transformação de envelope temporal, após realizar a filtragem de previsão linear na direção de frequência no componente de alta frequência no domínio de frequência gerado pelo meio de geração de alta frequência, preferencialmente ajusta a força de um componente de alta frequência obtido como resultado da filtragem de previsão linear em um valor equivalente àquele antes da filtragem de previsão linear.

No dispositivo de decodificação de voz da presente invenção, o meio de transformação de envelope temporal, após realizar a filtragem de previsão linear na direção de frequência no componente de alta frequência no domínio de frequência gerado pelo meio de geração de alta frequência, preferencialmente ajusta a força em uma certa faixa de frequência de um componente de alta frequência obtido como resultado da filtragem de previsão linear em um valor equivalente àquele antes da filtragem de previsão linear.

No dispositivo de decodificação de voz da presente invenção, as informações suplementares de envelope temporal são preferencialmente uma razão de um valor mínimo para um valor médio das informações de envelope temporal ajustadas.

No dispositivo de decodificação de voz da presente invenção, o meio de transformação de envelope temporal, após controlar o ganho do envelope temporal ajustado de forma que a força do componente de alta frequência no domínio de frequência em um segmento de tempo de envelope de SBR seja equivalente antes e depois da transformação do envelope temporal, preferencialmente transforma um envelope temporal do componente

de alta frequência multiplicando o envelope temporal cujo ganho é controlado pelo componente de alta frequência no domínio de frequência.

No dispositivo de decodificação de voz da presente invenção, o meio de análise de envelope temporal de baixa frequência preferencialmente obtém força de cada amostra de sub-banda de QMF do componente de baixa frequência convertido ao domínio de frequência pelo meio de conversão de frequência, e obtém as informações de envelope temporal representadas como um coeficiente de ganho a ser multiplicado por cada uma das amostras de sub-banda de QMF, normalizando a força de cada uma das amostras de sub-banda de QMF com uso de uma força média em um segmento de tempo de envelope de SBR.

Um dispositivo de decodificação de voz da presente invenção é um dispositivo de decodificação de voz para decodificar um sinal de voz codificado e inclui: meio de decodificação central para obter um componente de baixa frequência decodificando um fluxo de bit recebido de fora do dispositivo de decodificação que inclui o sinal de voz codificado; meio de conversão de frequência para converter o componente de baixa frequência obtido pelo meio de decodificação central em um domínio de frequência; meio de geração de alta frequência para gerar um componente de alta frequência copiando o componente de baixa frequência convertido no domínio de frequência pelo meio de conversão de frequência de uma banda de baixa frequência para uma banda de alta frequência; meio de análise de envelope temporal de baixa frequência para analisar o componente de baixa frequência convertido no domínio de frequência pelo meio de conversão de frequência para obter informações de envelope temporal; meio de geração de informações suplementares de envelope temporal para analisar o fluxo de bit para gerar informações suplementares de envelope temporal; meio de ajuste de envelope temporal para ajustar as informações de envelope temporal obtidas pelo meio de análise de envelope temporal de baixa frequência com uso das informações suplementares de envelope temporal; e meio de transformação de envelope temporal para transformar um envelope temporal do componente de alta frequência gerado pelo meio de geração de alta fre-

quência com uso das informações de envelope temporal ajustadas pelo meio de ajuste de envelope temporal.

É preferível que o dispositivo de decodificação de voz da presente invenção inclua um meio de ajuste de alta frequência primário e um meio de ajuste de alta frequência secundário, ambos correspondendo ao meio de ajuste de alta frequência, o meio de ajuste de alta frequência primário pode executar um processo que inclui uma parte de um processo que corresponde ao meio de ajuste de alta frequência, o meio de transformação de envelope temporal pode transformar um envelope temporal de um sinal de sávida do meio de ajuste de alta frequência primário, o meio de ajuste de alta frequência secundário pode executar um processo não executado pelo meio de ajuste de alta frequência primário dentro os processos que correspondem ao meio de ajuste de alta frequência em um sinal de sávida do meio de transformação de envelope temporal, e o meio de ajuste de alta frequência secundário pode ser um processo de adição de uma onda de seno durante a decodificação de SBR.

Efeitos Vantajosos da Invenção

De acordo com a presente invenção, a ocorrência de pré-eco e pós-eco pode ser reduzida e a qualidade subjetiva de um sinal decodificado pode ser melhorada sem aumentar significativamente a taxa de bit na técnica de extensão de banda no domínio de frequência representado pela SBR.

Breve Descrição dos Desenhos

A figura 1 é um diagrama que ilustra um dispositivo de codificação de voz de acordo com uma primeira modalidade;

A figura 2 é um fluxograma para descrever uma operação do dispositivo de codificação de voz de acordo com a primeira modalidade;

A figura 3 é um diagrama que ilustra um dispositivo de decodificação de voz de acordo com a primeira modalidade;

A figura 4 é um fluxograma para descrever uma operação do dispositivo de decodificação de voz de acordo com a primeira modalidade;

A figura 5 é um diagrama que ilustra um dispositivo de codificação de voz de acordo com uma primeira modificação da primeira modalidade;

5 A figura 6 é um diagrama que ilustra um dispositivo de codificação de voz de acordo com uma segunda modalidade;

A figura 7 é um fluxograma para descrever uma operação do dispositivo de codificação de voz de acordo com a segunda modalidade;

A figura 8 é um diagrama que ilustra um dispositivo de decodificação de voz de acordo com a segunda modalidade;

10 A figura 9 é um fluxograma para descrever uma operação do dispositivo de decodificação de voz de acordo com a segunda modalidade;

A figura 10 é um diagrama que ilustra um dispositivo de codificação de voz de acordo com uma terceira modalidade;

15 A figura 11 é um fluxograma para descrever uma operação do dispositivo de codificação de voz de acordo com a terceira modalidade;

A figura 12 é um diagrama que ilustra um dispositivo de decodificação de voz de acordo com a terceira modalidade;

A figura 13 é um fluxograma para descrever uma operação do dispositivo de decodificação de voz de acordo com a terceira modalidade;

20 A figura 14 é um diagrama que ilustra um dispositivo de decodificação de voz de acordo com uma quarta modalidade;

A figura 15 é um diagrama que ilustra um dispositivo de decodificação de voz de acordo com uma modificação da quarta modalidade;

25 A figura 16 é um diagrama que ilustra um dispositivo de decodificação de voz de acordo com outra modificação da quarta modalidade;

A figura 17 é um fluxograma para descrever uma operação do dispositivo de decodificação de voz de acordo com a outra modificação da quarta modalidade;

30 A figura 18 é um diagrama que ilustra um dispositivo de decodificação de voz de acordo com outra modificação da primeira modalidade;

A figura 19 é um fluxograma para descrever uma operação do dispositivo de decodificação de voz de acordo com a outra modificação da primeira modalidade;

5 A figura 20 é um diagrama que ilustra um dispositivo de decodificação de voz de acordo com outra modificação da primeira modalidade;

A figura 21 é um fluxograma para descrever uma operação do dispositivo de decodificação de voz de acordo com a outra modificação da primeira modalidade.

10 A figura 22 é um diagrama que ilustra um dispositivo de decodificação de voz de acordo com uma modificação da segunda modalidade;

A figura 23 é um fluxograma para descrever uma operação do dispositivo de decodificação de voz de acordo com a outra modificação da segunda modalidade;

15 A figura 24 é um diagrama que ilustra um dispositivo de decodificação de voz de acordo com outra modificação da segunda modalidade;

A figura 25 é um fluxograma para descrever uma operação do dispositivo de decodificação de voz de acordo com a outra modificação da segunda modalidade;

20 A figura 26 é um diagrama que ilustra um dispositivo de decodificação de voz de acordo com outra modificação da quarta modalidade;

A figura 27 é um fluxograma para descrever uma operação do dispositivo de decodificação de voz de acordo com a outra modificação da quarta modalidade;

25 A figura 28 é um diagrama de um dispositivo de decodificação de voz de acordo com outra modificação da quarta modalidade;

A figura 29 é um fluxograma para descrever uma operação do dispositivo de decodificação de voz de acordo com a outra modificação da quarta modalidade;

30 A figura 30 é um diagrama que ilustra um dispositivo de decodificação de voz de acordo com outra modificação da quarta modalidade;

A figura 31 é um diagrama que ilustra um dispositivo de decodificação de voz de acordo com outra modificação da quarta modalidade;

A figura 32 é um fluxograma para descrever uma operação do dispositivo de decodificação de voz de acordo com a outra modificação da quarta modalidade;

5 A figura 33 é um diagrama que ilustra um dispositivo de decodificação de voz de acordo com outra modificação da quarta modalidade;

A figura 34 é um fluxograma para descrever uma operação do dispositivo de decodificação de voz de acordo com a outra modificação da quarta modalidade;

10 A figura 35 é um diagrama que ilustra um dispositivo de decodificação de voz de acordo com outra modificação da quarta modalidade;

A figura 36 é um fluxograma para descrever uma operação do dispositivo de decodificação de voz de acordo com a outra modificação da quarta modalidade;

15 A figura 37 é um diagrama que ilustra um dispositivo de decodificação de voz de acordo com outra modificação da quarta modalidade;

A figura 38 é um diagrama que ilustra um dispositivo de decodificação de voz de acordo com outra modificação da quarta modalidade;

20 A figura 39 é um fluxograma para descrever uma operação do dispositivo de decodificação de voz de acordo com a outra modificação da quarta modalidade;

A figura 40 é um diagrama que ilustra um dispositivo de decodificação de voz de acordo com outra modificação da quarta modalidade;

25 A figura 41 é um fluxograma para descrever uma operação do dispositivo de decodificação de voz de acordo com a outra modificação da quarta modalidade;

A figura 42 é um diagrama que ilustra um dispositivo de decodificação de voz de acordo com outra modificação da quarta modalidade;

30 A figura 43 é um fluxograma para descrever uma operação do dispositivo de decodificação de voz de acordo com a outra modificação da quarta modalidade;

A figura 44 é um diagrama que ilustra um dispositivo de codificação de voz de acordo com outra modificação da primeira modalidade;

A figura 45 é um diagrama que ilustra um dispositivo de codificação de voz de acordo com still outra modificação da primeira modalidade;

A figura 46 é um diagrama que ilustra um dispositivo de codificação de voz de acordo com uma modificação da segunda modalidade;

5 A figura 47 é um diagrama que ilustra um dispositivo de codificação de voz de acordo com outra modificação da segunda modalidade;

A figura 48 é um diagrama que ilustra um dispositivo de codificação de voz de acordo com a quarta modalidade;

10 A figura 49 é um diagrama que ilustra um dispositivo de codificação de voz de acordo com uma modificação da quarta modalidade; e

A figura 50 é um diagrama que ilustra um dispositivo de codificação de voz de acordo com outra modificação da quarta modalidade.

Descrição das Modalidades

15 As modalidades preferenciais de acordo com a presente invenção são descritas abaixo em detalhe com referência aos desenhos anexos. Na descrição dos desenhos, os elementos que são os mesmos são indicados com os mesmos símbolos de referência e a descrição duplicada dos mesmos é omitida, se aplicável.

(Primeira modalidade)

20 A figura 1 é um diagrama que ilustra um dispositivo de codificação de voz 11 de acordo com a primeira modalidade. O dispositivo de codificação de voz 11 inclui fisicamente uma CPU, uma ROM, uma RAM, um dispositivo de comunicação e similares, que não são ilustrados, e a CPU controla integralmente o dispositivo de codificação de voz 11 carregando e executando um programa de computador predeterminado (tal como um programa de computador para realizar os processos ilustrados no fluxograma da
25 figura 2) armazenado em uma memória interna do dispositivo de codificação de voz 11 tal como a ROM na RAM. O dispositivo de comunicação do dispositivo de codificação de voz 11 recebe um sinal de voz a ser codificado de
30 fora do dispositivo de codificação de voz 11, e emite um fluxo de bit multiplexado codificado para fora do dispositivo de codificação de voz 11.

O dispositivo de codificação de voz 11 funcionalmente inclui uma unidade de conversão de frequência 1a (meio de conversão de frequência), uma unidade de conversão inversa de frequência 1b, a unidade de codificação de codec central 1c (meio de codificação central), uma unidade de codificação de SBR 1d, uma unidade de análise de previsão linear 1e (meio de cálculo de informações suplementares de envelope temporal), uma unidade de cálculo de parâmetro de resistência de filtro 1f (meio de cálculo de informações suplementares de envelope temporal), e uma unidade de multiplexação de fluxo de bit 1g (meio de multiplexação de fluxo de bit). A unidade de conversão de frequência 1a à unidade de multiplexação de fluxo de bit 1g do dispositivo de codificação de voz 11 ilustradas na figura 1 são funções realizadas quando a CPU do dispositivo de codificação de voz 11 executa o programa de computador armazenado na memória interna do dispositivo de codificação de voz 11. A CPU do dispositivo de codificação de voz 11 sequencialmente executa processos (processos da Etapa Sa1 à Etapa Sa7) ilustrados no fluxograma da figura 2, executando o programa de computador (ou usando a unidade de conversão de frequência 1a à unidade de multiplexação de fluxo de bit 1g ilustradas na figura 1). Vários tipos de dados necessários para executar o programa de computador e vários tipos de dados gerados para executar o programa de computador são todos armazenados na memória interna tal como a ROM e a RAM do dispositivo de codificação de voz 11.

A unidade de conversão de frequência 1a analisa um sinal de entrada recebido de fora do dispositivo de codificação de voz 11 por meio do dispositivo de comunicação do dispositivo de codificação de voz 11 com uso de um banco de filtro de QMF de múltiplas divisões para obter um sinal $q(k, r)$ em um domínio de QMF (processo na Etapa Sa1). Nota-se que k ($0 \leq k \leq 63$) é um índice em uma direção de frequência, e r é um índice que indica uma abertura de tempo. A unidade de conversão inversa de frequência 1b sintetiza uma metade dos coeficientes no lado de baixa frequência do domínio de QMF obtidos pela unidade de conversão de frequência 1a com uso do banco de filtro de QMF para obter um sinal de domínio de tempo de resolução re-

duzida que inclui somente componentes de baixa frequência do sinal de entrada (processo na Etapa Sa2). A unidade de codificação de codec central 1c codificou o sinal de domínio de resolução reduzida para obter um fluxo de bit codificado (processo na Etapa Sa3). A codificação realizada pela unidade de codificação de codec central 1c pode estar baseada em um método de codificação de voz representado por um método CELP ou pode estar baseada em uma codificação de transformação representada por AAC ou uma codificação de som tal como um método de TCX (Excitação Codificada de Transformação).

10 A unidade de codificação de SBR 1d recebe o sinal no domínio de QMF a partir da unidade de conversão de frequência 1a, e realiza a codificação de SBR com base na análise da força, alteração de sinal, tonalidade e similares dos componentes de alta frequência para obter informações suplementares de SBR (processo na Etapa Sa4). O método de análise de QMF na unidade de conversão de frequência 1a e o método de codificação de SBR na unidade de codificação de SBR 1d são descritos em detalhe, por exemplo, em uma Literatura "3GPP TS 26.404: Enhanced aacPlus encoder SBR part".

20 A unidade de análise de previsão linear 1e recebe o sinal no domínio de QMF a partir da unidade de conversão de frequência 1a, e realiza a análise de previsão linear na direção de frequência nos componentes de alta frequência do sinal para obter coeficientes de previsão linear de alta frequência $a_H(n, r)$ ($1 \leq n \leq N$) (processo na Etapa Sa5). Nota-se que N é uma ordem de previsão linear. O índice r é um índice em uma direção temporal para uma sub-amostra dos sinais no domínio de QMF. Um método de covariância ou um método de auto-correlação podem ser usados para a análise de previsão linear de sinal. A análise de previsão linear para obter $a_H(n, r)$ é realizada nos componentes de alta frequência que satisfazem $k_x < k \leq 63$ em $q(k, r)$. Nota-se que k_x é um índice de frequência que corresponde a uma frequência de limite superior da banda de frequência codificada pela unidade de codificação de codec central 1c. A unidade de análise de previsão linear 1e pode também realizar a análise de previsão linear em componentes de

baixa frequência diferentes daqueles analisados quando $a_H(n, r)$ são obtidos para obter coeficientes de previsão linear de baixa frequência $a_L(n, r)$ diferentes de $a_H(n, r)$ (coeficientes de previsão linear de acordo com tais componentes de baixa frequência correspondem às informações de envelope temporal, e é o mesmo na primeira modalidade conforme abaixo). A análise de previsão linear para obter $a_L(n, r)$ é realizada em componentes de baixa frequência que satisfazem $0 \leq k < k_x$. A análise de previsão linear pode ser também realizada em uma parte banda de frequência incluída em uma seção de $0 \leq k < k_x$.

- 10 A unidade de cálculo de parâmetro de resistência de filtro 1f, por exemplo, utiliza os coeficientes de previsão linear obtidos pela unidade de análise de previsão linear 1e para calcular um parâmetro de resistência de filtro (o parâmetro de resistência de filtro corresponde às informações suplementares de envelope temporal e é o mesmo na primeira modalidade conforme mostrado abaixo) (processo na Etapa Sa6). Um ganho de previsão $G_H(r)$ é o primeiro calculado a partir de $a_H(n, r)$. O método para calcular o ganho de previsão é, por exemplo, descrito em detalhe em "Speech Coding, Takehiro Moriya, The Institute of Electronics, Information and Communication Engineers". Se $a_L(n, r)$ foi calculado, um ganho de previsão $G_L(r)$ é calculado
- 15 semelhante. O parâmetro de resistência de filtro $K(r)$ é um parâmetro que aumenta conforme $G_H(r)$ é aumentado, e, por exemplo, pode ser obtido de acordo com a seguinte expressão (1). Aqui, $\max(a, b)$ indica o valor máximo de a e b , e $\min(a, b)$ indica o valor mínimo de a e b .
- 20

$$K(r) = \max(0, \min(1, G_H(r) - 1)) \quad \text{---(1)}$$

- 25 Se $G_L(r)$ foi calculado, $K(r)$ pode ser obtido como um parâmetro que aumenta conforme $G_H(r)$ é aumentado, e diminui conforme $G_L(r)$ é aumentado. Neste caso, por exemplo, K pode ser obtido de acordo com a seguinte expressão (2).

$$K(r) = \max(0, \min(1, G_H(r)/G_L(r) - 1)) \quad \text{---(2)}$$

- 30 $K(r)$ é um parâmetro que indica a resistência para ajustar o envelope temporal dos componentes de alta frequência durante a decodificação de SBR. Um valor do ganho de previsão com relação aos coeficientes de

previsão linear na direção de frequência é aumentado conforme a variação do envelope temporal de um sinal na seção analisada se torna agudo. $K(r)$ é um parâmetro para instruir um decodificador a reforçar o processo para alterar intensamente o envelope temporal dos componentes de alta frequência gerados por SBR, com o aumento de seu valor. $K(r)$ pode ser também um parâmetro para instruir um decodificador (tal como um dispositivo de decodificação de voz 21) a enfraquecer o processo para variar intensamente o envelope temporal dos componentes de alta frequência gerados por SBR, com a diminuição de seu valor, ou pode incluir um valor para não executar o processo para variar intensamente o envelope temporal. Ao invés de transmitir $K(r)$ para cada abertura de tempo, $K(r)$ que represente uma pluralidade de aberturas de tempo pode ser transmitido. Para determinar as seções de abertura de tempo que compartilham o mesmo valor de $K(r)$, é preferível usar as informações de limite de tempo de envelope de SBR incluídas nas informações suplementares de SBR.

$K(r)$ é transmitido à unidade de multiplexação de fluxo de bit 1g após ser quantizado. É preferível calcular $K(r)$ que representa a pluralidade de aberturas de tempo, por exemplo, tomando uma média de $K(r)$ de uma pluralidade de aberturas de tempo r antes de a quantização ser realizada. Para transmitir $K(r)$ que representa a pluralidade de aberturas de tempo, $K(r)$ pode ser também obtido a partir do resultado da análise de todas as seções formadas da pluralidade de aberturas de tempo, ao invés de independentemente calcular $K(r)$ a partir do resultado da análise de cada abertura de tempo tal como a expressão (2). Neste caso, $K(r)$ pode ser calculado, por exemplo, de acordo com a seguinte expressão (3). Aqui, $\text{mean}(\cdot)$ indica um valor médio nas seções de abertura de tempo representadas por $K(r)$.

$$K(r) = \max(0, \min(1, \text{mean}(G_H(r))/\text{mean}(G_L(r)) - 1)))$$

---(3)

$K(r)$ pode ser exclusivamente transmitido com informações de modo de filtro inverso incluídas nas informações suplementares de SBR descritas em "ISO/IEC 14496-3 subpart 4 General Audio Coding". Em outras palavras, $K(r)$ não é transmitido para a abertura de tempo para transmitir as

informações de modo de filtro inverso nas informações suplementares de SBR, e as informações de modo de filtro inverso (bs#invf#mode em "ISO/IEC 14496-3 subpart 4 General Audio Coding") nas informações suplementares de SBR não precisam ser transmitidas para a abertura de tempo para transmitir K(r). As informações que indicam se tanto K(r) ou as informações de modo de filtro inverso incluídas nas informações suplementares de SBR são transmitidas podem ser também adicionadas. K(r) e as informações de modo de filtro inverso incluídas nas informações suplementares de SBR podem ser combinadas para funcionarem como informações de vetor, e realizarem codificação de entropia no vetor. Neste caso, a combinação de K(r) e o valor das informações de modo de filtro inverso incluídas nas informações suplementares de SBR podem ser restringidos.

A unidade de multiplexação de fluxo de bit 1g multiplexa o fluxo de bit codificado calculado pela unidade de codificação de codec central 1c, as informações suplementares de SBR calculadas pela unidade de codificação de SBR 1d, e K(r) calculado pela unidade de cálculo de parâmetro de resistência de filtro 1f, e emite um fluxo de bit multiplexado (fluxo de bit multiplexado codificado) através do dispositivo de comunicação do dispositivo de codificação de voz 11 (processo na Etapa Sa7).

A figura 3 é um diagrama que ilustra um dispositivo de decodificação de voz 21 de acordo com a primeira modalidade. O dispositivo de decodificação de voz 21 fisicamente inclui uma CPU, uma ROM, uma RAM, um dispositivo de comunicação, e similares, que não são ilustrados, e a CPU integralmente controla o dispositivo de decodificação de voz 21 carregando e executando um programa de computador predeterminado (tal como um programa de computador para realizar os processos ilustrados no fluxograma da figura 4) armazenado em uma memória interna do dispositivo de decodificação de voz 21 tal como a ROM na RAM. O dispositivo de comunicação do dispositivo de decodificação de voz 21 recebe a saída de fluxo de bit multiplexado a partir do dispositivo de codificação de voz 11, um dispositivo de codificação de voz 11a de uma modificação 1, que será descrito adiante, ou um dispositivo de codificação de voz de uma modificação 2, que será descri-

to adiante, e emite um sinal de voz decodificado para fora do dispositivo de decodificação de voz 21. O dispositivo de decodificação de voz 21, conforme ilustrado na figura 3, funcionalmente inclui uma unidade de separação de fluxo de bit 2a (meio de separação de fluxo de bit), uma unidade de decodifi-
5 cação de codec central 2b (meio de decodificação central), uma unidade de conversão de frequência 2c (meio de conversão de frequência), uma unidade de análise de previsão linear de baixa frequência 2d (meio de análise de envelope temporal de baixa frequência), uma unidade de detecção de sinal 2e, uma unidade de ajuste de resistência de filtro 2f (meio de ajuste de enve-
10 lope temporal), uma unidade de geração de alta frequência 2g (meio de geração de alta frequência), uma unidade de análise de previsão linear de alta frequência 2h, uma unidade de filtro inverso de previsão linear 2i, uma unidade de ajuste de alta frequência 2j (meio de ajuste de alta frequência), uma unidade de filtro de previsão linear 2k (meio de transformação de envelope
15 temporal), uma unidade de adição de coeficiente 2m e uma unidade de conversão inversa de frequência 2n. A unidade de separação de fluxo de bit 2a a uma unidade de transformação inversa de frequência 2n do dispositivo de decodificação de voz 21 ilustrado na figura 3 são funções realizadas quando a CPU do dispositivo de decodificação de voz 21 executa o programa de
20 computador armazenado na memória interna do dispositivo de decodificação de voz 21. A CPU do dispositivo de decodificação de voz 21 sequencialmente executa os processos (processos da Etapa Sb1 à Etapa Sb11) ilustrados no fluxograma da figura 4, executando o programa de computador (ou com uso da unidade de separação de fluxo de bit 2a à unidade de transformação
25 inversa de frequência 2n ilustrada na figura 3). Vários tipos de dados necessários para executar o programa de computador e vários tipos de dados gerados executando-se o programa de computador são todos armazenados na memória interna tal como a ROM e a RAM do dispositivo de decodificação de voz 21.

30 A unidade de separação de fluxo de bit 2a separa o fluxo de bit multiplexado fornecido através do dispositivo de comunicação do dispositivo de decodificação de voz 21 em um parâmetro de resistência de filtro, infor-

mações suplementares de SBR e o fluxo de bit codificado. A unidade de decodificação de codec central 2b decodifica o fluxo de bit codificado recebido a partir da unidade de separação de fluxo de bit 2a para obter um sinal decodificado que inclui somente os componentes de baixa frequência (processo na Etapa Sb1). Neste momento, o método de decodificação pode ser baseado no método de codificação de voz representado pelo método CELP, ou pode ser baseado em decodificação de som tal como o método AAC ou o TCX (Excitação Codificada de Transformação).

A unidade de conversão de frequência 2c analisa o sinal decodificado recebido a partir da unidade de decodificação de codec central 2b com uso do banco de filtro de QMF de múltiplas divisões para obter um sinal $q_{dec}(k, r)$ no domínio de QMF (processo na Etapa Sb2). Nota-se que k ($0 \leq k \leq 63$) é um índice na direção de frequência, e r é um índice que indica um índice para a sub-amostra do sinal no domínio de QMF na direção temporal.

A unidade de análise de previsão linear de baixa frequência 2d realiza a análise de previsão linear na direção de frequência em $q_{dec}(k, r)$ de cada abertura de tempo r , obtida a partir da unidade de conversão de frequência 2c, para obter coeficientes de previsão linear de baixa frequência $a_{dec}(n, r)$ (processo na Etapa Sb3). A análise de previsão linear é realizada para uma faixa de $0 \leq k < k_x$ que corresponde a uma banda de sinal do sinal decodificado obtido a partir da unidade de decodificação de codec central 2b. A análise de previsão linear pode ser realizada em uma parte da banda de frequência incluída na seção de $0 \leq k < k_x$.

A unidade de detecção de sinal 2e detecta a variação temporal do sinal no domínio de QMF recebido a partir da unidade de conversão de frequência 2c, e emite-o como um resultado de detecção $T(r)$. A alteração de sinal pode ser detectada, por exemplo, com uso do método descrito abaixo.

1. A força de período curto $p(r)$ de um sinal na abertura de tempo r é obtida de acordo com a seguinte expressão (4).

$$p(r) = \sum_{k=0}^{63} |q_{dec}(k, r)|^2 \quad \text{---(4)}$$

2. Um envelope $p_{env}(r)$ obtido suavizando-se $p(r)$ é obtido de acordo com a seguinte expressão (5). Note-se que α é uma constante que satisfaz $0 < \alpha < 1$.

$$p_{env}(r) = \alpha \cdot p_{env}(r-1) + (1-\alpha) \cdot p(r) \quad \text{---(5)}$$

3. $T(r)$ é obtido de acordo com a seguinte expressão (6)
5 com uso de $p(r)$ e $p_{env}(r)$, em que β é uma constante.

$$T(r) = \max(1, p(r)/(\beta \cdot p_{env}(r))) \quad \text{---(6)}$$

Os métodos descritos acima são simples exemplos para detectar a alteração de sinal baseada na alteração na força, e a alteração de sinal pode ser detectada com uso de outros métodos mais sofisticados. Em adição, a unidade de detecção de sinal 2e pode ser omitida.

10 A unidade de ajuste de resistência de filtro 2f ajusta a resistência de filtro com relação a $a_{dec}(n, r)$ obtido a partir da unidade de análise de previsão linear de baixa frequência 2d para obter um coeficiente de previsão linear ajustado $a_{adj}(n, r)$, (processo na Etapa Sb4). A resistência de filtro é ajustada, por exemplo, de acordo com a seguinte expressão (7), com uso de
15 um parâmetro de resistência de filtro K recebido através da unidade de separação de fluxo de bit 2a.

$$a_{adj}(n, r) = a_{dec}(n, r) \cdot K(r)^n \quad (1 \leq n \leq N) \quad \text{---(7)}$$

Se uma saída $T(r)$ é obtida a partir da unidade de detecção de sinal 2e, a resistência pode ser ajustada de acordo com a seguinte expressão (8).

$$a_{adj}(n, r) = a_{dec}(n, r) \cdot (K(r) \cdot T(r))^n \quad (1 \leq n \leq N) \quad \text{---(8)}$$

20 A unidade de geração de alta frequência 2g copia o sinal no domínio de QMF obtido a partir da unidade de conversão de frequência 2c a partir da banda de baixa frequência para a banda de alta frequência para gerar um sinal $q_{exp}(k, r)$ no domínio de QMF dos componentes de alta frequência (processo na Etapa Sb5). Os componentes de alta frequência são
25 gerados de acordo com o método de geração HF em SBR em "MPEG4 AAC" ("ISO/IEC 14496-3 subpart 4 General Audio Coding").

A unidade de análise de previsão linear de alta frequência 2h realiza a análise de previsão linear na direção de frequência em $q_{exp}(k, r)$ de cada uma das aberturas de tempo r geradas pela unidade de geração de alta frequência 2g para obter os coeficientes de previsão linear de alta frequência $a_{exp}(n, r)$ (processo na Etapa Sb6). A análise de previsão linear é realizada para uma faixa de $k_x \leq k \leq 63$ que corresponde aos componentes de alta frequência gerados pela unidade de geração de alta frequência 2g.

A unidade de filtro inverso de previsão linear 2i realiza a filtragem inversa de previsão linear na direção de frequência em um sinal no domínio de QMF da banda de alta frequência gerada pela unidade de geração de alta frequência 2g, com uso de $a_{exp}(n, r)$ como coeficientes (processo na Etapa Sb7). A função de transferência do filtro inverso de previsão linear pode ser expressa como a seguinte expressão (9)

$$f(z) = 1 + \sum_{n=1}^N a_{exp}(n, r) z^{-n} \quad \text{---(9)}$$

A filtragem inversa de previsão linear pode ser realizada a partir do coeficiente no lado de baixa frequência em direção ao coeficiente no lado de alta frequência, ou pode ser realizada na direção oposta. A filtragem inversa de previsão linear é um processo para temporariamente suavizar o envelope temporal dos componentes de alta frequência, antes da transformação do envelope temporal ser realizada no estágio subsequente, e a unidade de filtro inverso de previsão linear 2i pode ser omitida. É também possível realizar a análise de previsão linear e filtragem inversa em saídas da unidade de ajuste de alta frequência 2j, que será descrita adiante, pela unidade de análise de previsão linear de alta frequência 2ha e pela unidade de filtro inverso de previsão linear 2i, ao invés de realizar a análise de previsão linear e filtragem inversa nos componentes de alta frequência das saídas da unidade de geração de alta frequência 2g. Os coeficientes de previsão linear usados para a filtragem inversa de previsão linear podem ser também $a_{dec}(n, r)$ ou $a_{adj}(n, r)$, ao invés de $a_{exp}(n, r)$. Os coeficientes de previsão linear usados para a filtragem inversa de previsão linear podem ser também coeficientes de previsão linear $a_{exp,adj}(n, r)$ obtidos realizando-se o ajuste de resistên-

cia de filtro em $a_{exp}(n, r)$. O ajuste de resistência é realizado de acordo com a seguinte expressão (10), similar àquela em que $a_{adj}(n, r)$ é obtido.

$$a_{exp,adj}(n, r) = a_{exp}(n, r) \cdot K(r)^n \quad (1 \leq n \leq N) \quad \text{---(10)}$$

A unidade de ajuste de alta frequência 2j ajusta as características de frequência e tonalidade dos componentes de alta frequência de uma saída da unidade de filtro inverso de previsão linear 2i (processo na Etapa Sb8). O ajuste é realizado de acordo com as informações suplementares de SBR recebidas a partir da unidade de separação de fluxo de bit 2a. O processamento pela unidade de ajuste de alta frequência 2j é realizado de acordo com etapa de "ajuste HF" em SBR em "MPEG4 AAC", e é ajustado realizando-se a filtragem inversa de previsão linear na direção temporal, o ajuste de ganho, e a superposição de ruído no sinal do domínio de QMF da banda de alta frequência. Os detalhes dos processos nas etapas descritas acima são descritos em "ISO/IEC 14496-3 subpart 4 General Audio Coding". Conforme descrito acima, a unidade de conversão de frequência 2c, a unidade de geração de alta frequência 2g, e a unidade de ajuste de alta frequência 2j todas operam de acordo com o decodificador SBR em "MPEG4 AAC" definido em "ISO/IEC 14496-3".

A unidade de filtro de previsão linear 2k realiza filtragem de síntese de previsão linear na direção de frequência em componentes de alta frequência $q_{adj}(n, r)$ de um sinal na saída de domínio de QMF a partir da unidade de ajuste de alta frequência 2j, com uso de $a_{adj}(n, r)$ obtido a partir da unidade de ajuste de resistência de filtro 2f (processo na Etapa Sb9). A função de transferência na filtragem de síntese de previsão linear pode ser expressa como a seguinte expressão (11).

$$g(z) = \frac{1}{1 + \sum_{n=1}^N a_{adj}(n, r) z^{-n}} \quad \text{---(11)}$$

Realizando-se a filtragem de síntese de previsão linear, a unidade de filtro de previsão linear 2k transforma o envelope temporal dos componentes de alta frequência gerados com base em SBR.

A unidade de adição de coeficiente $2m$ adiciona um sinal no domínio de QMF que inclui a saída dos componentes de baixa frequência da unidade de conversão de frequência $2c$ e um sinal no domínio de QMF que inclui a saída dos componentes de alta frequência da unidade de filtro de previsão linear $2k$, e emite um sinal no domínio de QMF que inclui ambos os componentes de baixa frequência e os componentes de alta frequência (processo na Etapa Sb10).

A unidade de conversão inversa de frequência $2n$ processa o sinal no domínio de QMF obtido a partir da unidade de adição de coeficiente $2m$ com uso de um banco de filtro de síntese de QMF. Desta forma, um domínio de tempo decodificou um sinal de voz que inclui ambos os componentes de baixa frequência obtidos sendo decodificados pelo codec central e os componentes de alta frequência gerados por SBR e cujo envelope temporal é transformado pelo filtro de previsão linear é obtido, e o sinal obtido de voz é emitido para fora do dispositivo de decodificação de voz 21 através do dispositivo de comunicação interno (processo na Etapa Sb11). Se $K(r)$ e as informações de modo de filtro inverso das informações suplementares de SBR descritas em "ISO/IEC 14496-3 subpart 4 General Audio Coding" forem exclusivamente transmitidas, a unidade de conversão inversa de frequência $2n$ pode gerar informações de modo de filtro inverso das informações suplementares de SBR para uma abertura de tempo para a qual $K(r)$ é transmitido, mas as informações de modo de filtro inverso das informações suplementares de SBR não são transmitidas, com uso das informações de modo de filtro inverso das informações suplementares de SBR com relação a ao menos uma abertura de tempo da abertura de tempos entre a abertura de tempos antes e após a abertura de tempo. É também possível definir as informações de modo de filtro inverso das informações suplementares de SBR da abertura de tempo em um modo predeterminado antecipadamente. A unidade de conversão inversa de frequência $2n$ pode gerar $K(r)$ para uma abertura de tempo para a qual os dados de filtro inverso das informações suplementares de SBR são transmitidos, mas $K(r)$ não é transmitido, com uso de $K(r)$ para ao menos uma abertura de tempo entre a abertura de tempos an-

tes e após a abertura de tempo. É também possível definir $K(r)$ da abertura de tempo em um valor predeterminado antecipadamente. A unidade de conversão inversa de frequência $2n$ pode determinar também se as informações transmitidas são $K(r)$ ou as informações de modo de filtro inverso das informações suplementares de SBR, com base nas informações que indicam se $K(r)$ ou as informações de modo de filtro inverso das informações suplementares de SBR são transmitidos.

(Modificação 1 da Primeira Modalidade)

A figura 5 é um diagrama que ilustra a modificação (dispositivo de codificação de voz 11a) do dispositivo de codificação de voz de acordo com a primeira modalidade. O dispositivo de codificação de voz 11a fisicamente inclui uma CPU, uma ROM, uma RAM, um dispositivo de comunicação, e similares, que não são ilustrados, e a CPU controla integralmente o dispositivo de codificação de voz 11a carregando e executando um programa de computador predeterminado armazenado em uma memória interna do dispositivo de codificação de voz 11a tal como a ROM na RAM. O dispositivo de comunicação do dispositivo de codificação de voz 11a recebe um sinal de voz a ser codificado de fora do dispositivo de codificação de voz 11a, e emite um fluxo de bit multiplexado codificado para fora do dispositivo de codificação de voz 11a.

O dispositivo de codificação de voz 11a, conforme ilustrado na figura 5, funcionalmente inclui uma unidade de conversão inversa de alta frequência 1h, uma unidade de cálculo de força de período curto 1i (meio de cálculo de informações suplementares de envelope temporal), uma unidade de cálculo de parâmetro de resistência de filtro 1f1 (meio de cálculo de informações suplementares de envelope temporal), e uma unidade de multiplexação de fluxo de bit 1g1 (meio de multiplexação de fluxo de bit), ao invés da unidade de análise de previsão linear 1e, da unidade de cálculo de parâmetro de resistência de filtro 1f, e da unidade de multiplexação de fluxo de bit 1g do dispositivo de codificação de voz 11. A unidade de multiplexação de fluxo de bit 1g1 tem a mesma função que a da unidade de multiplexação de fluxo de bit 1g. A unidade de conversão de frequência 1a à unidade de codi-

ficação de SBR 1d, a unidade de conversão inversa de alta frequência 1h, a unidade de cálculo de força de período curto 1i, a unidade de cálculo de parâmetro de resistência de filtro 1f1, e a unidade de multiplexação de fluxo de bit 1g1 do dispositivo de codificação de voz 11a ilustradas na figura 5 são

5 funções realizadas quando a CPU do dispositivo de codificação de voz 11a executa o programa de computador armazenado na memória interna do dispositivo de codificação de voz 11a. Vários tipos de dados necessários para executar o programa de computador e vários tipos de dados gerados executando-se o programa de computador são todos armazenados na memória

10 interna tal como a ROM e a RAM do dispositivo de codificação de voz 11a.

A unidade de conversão inversa de alta frequência 1h converte, no sinal no domínio de QMF obtido a partir da unidade de conversão de frequência 1a, coeficientes que correspondem aos componentes de baixa frequência codificado pela unidade de codificação de codec central 1c para "0",

15 para obter um sinal de domínio de tempo que inclui somente os componentes de alta frequência, e processar o sinal convertido com uso do banco de filtro de síntese de QMF. A unidade de cálculo de força de período curto 1i divide os componentes de alta frequência no domínio de tempo obtido a partir da unidade de conversão inversa de alta frequência 1h em seções curtas,

20 calcula a força, e calcula $p(r)$. Como um método alternativo, a força de período curto pode ser também calculada de acordo com a seguinte expressão (12) com uso do sinal no domínio de QMF.

$$p(r) = \sum_{k=0}^{63} |q(k, r)|^2 \quad \text{---(12)}$$

A unidade de cálculo de parâmetro de resistência de filtro 1f1 detecta a porção alterada de $p(r)$, e determina um valor de $K(r)$, de forma que

25 $K(r)$ seja aumentado com a alteração grande. O valor de $K(r)$, por exemplo, pode ser também calculado pelo mesmo método que aquele para calcular $T(r)$ pela unidade de detecção de sinal 2e do dispositivo de decodificação de voz 21. A alteração de sinal pode ser também detectada com uso de outros métodos mais sofisticados. A unidade de cálculo de parâmetro de resistência

30 de filtro 1f1 pode também obter força de período curto de cada um dos com-

ponentes de baixa frequência e dos componentes de alta frequência, obter alterações de sinal $Tr(r)$ e $Th(r)$ de cada um dos componentes de baixa frequência e dos componentes de alta frequência com uso do mesmo método que aquele para calcular $T(r)$ pela unidade de detecção de sinal 2e do dispositivo de decodificação de voz 21, e determina o valor de $K(r)$ com uso do mesmo. Neste caso, por exemplo, $K(r)$ pode ser obtido de acordo com a seguinte expressão (13), em que ε é uma constante tal como 3,0.

$$K(r) = \max(0, \varepsilon \cdot (Th(r) - Tr(r))) \quad \text{---(13)}$$

(Modificação 2 da Primeira Modalidade)

Um dispositivo de codificação de voz (não ilustrado) de uma modificação 2 da primeira modalidade inclui fisicamente uma CPU, uma ROM, uma RAM, um dispositivo de comunicação, e similares, que não são ilustrados, e a CPU controla integralmente o dispositivo de codificação de voz da modificação 2 carregando e executando um programa de computador predeterminado armazenado em uma memória interna do dispositivo de codificação de voz da modificação 2 tal como a ROM na RAM. O dispositivo de comunicação do dispositivo de codificação de voz da modificação 2 recebe um sinal de voz a ser codificado de fora do dispositivo de codificação de voz, e emite um fluxo de bit multiplexado codificado para fora do dispositivo de codificação de voz.

O dispositivo de codificação de voz da modificação 2 funcionalmente inclui uma unidade diferencial de codificação de coeficiente de previsão linear (meio de cálculo de informações suplementares de envelope temporal) e uma unidade de multiplexação de fluxo de bit (meio de multiplexação de fluxo de bit) que recebe uma saída a partir da unidade de codificação diferencial de coeficiente de previsão linear, que não são ilustradas, ao invés da unidade de cálculo de parâmetro de resistência de filtro 1f e da unidade de multiplexação de fluxo de bit 1g do dispositivo de codificação de voz 11. A unidade de conversão de frequência 1a à unidade de análise de previsão linear 1e, a unidade de codificação diferencial de coeficiente de previsão linear, e a unidade de multiplexação de fluxo de bit do dispositivo de codificação de voz da modificação 2 são funções realizadas quando a CPU do dis-

positivo de codificação de voz da modificação 2 executa o programa de computador armazenado na memória interna do dispositivo de codificação de voz da modificação 2. Vários tipos de dados necessários para executar o programa de computador e vários tipos de dados gerados executando-se o programa de computador são todos armazenados na memória interna tal como a ROM e a RAM do dispositivo de codificação de voz da modificação 2.

A unidade de codificação diferencial de coeficiente de previsão linear calcula valores diferenciais $a_D(n, r)$ do coeficiente de previsão linear de acordo com a seguinte expressão (14), com uso de $a_H(n, r)$ do sinal de entrada e $a_L(n, r)$ do sinal de entrada.

$$a_D(n, r) = a_H(n, r) - a_L(n, r) \quad (1 \leq n \leq N) \quad \text{---(14)}$$

A unidade de codificação diferencial de coeficiente de previsão linear quando quantiza $a_D(n, r)$, e transmite-os para a unidade de multiplexação de fluxo de bit (estrutura que corresponde à unidade de multiplexação de fluxo de bit 1g). A unidade de multiplexação de fluxo de bit multiplexa $a_D(n, r)$ no fluxo de bit ao invés de $K(r)$, e emite o fluxo de bit multiplexado para fora do dispositivo de codificação de voz através do dispositivo de comunicação interno.

Um dispositivo de decodificação de voz (não ilustrado) da modificação 2 da primeira modalidade inclui fisicamente uma CPU, uma ROM, uma RAM, um dispositivo de comunicação, e similares, que não são ilustrados, e a CPU controla integralmente o dispositivo de decodificação de voz da modificação 2 carregando e executando um programa de computador predeterminado armazenado em uma memória interna do dispositivo de decodificação de voz da modificação 2 tal como a ROM na RAM. O dispositivo de comunicação do dispositivo de decodificação de voz da modificação 2 recebe o fluxo de bit multiplexado codificado emitido a partir do dispositivo de codificação de voz 11, o dispositivo de codificação de voz 11a de acordo com a modificação 1, ou o dispositivo de codificação de voz de acordo com a modificação 2, e emite um sinal de voz decodificado para fora do dispositivo de decodificação de voz.

O dispositivo de decodificação de voz da modificação 2 funcionalmente inclui uma unidade de decodificação diferencial de coeficiente de previsão linear, que não é ilustrada, ao invés da unidade de ajuste de resistência de filtro 2f do dispositivo de decodificação de voz 21. A unidade de separação de fluxo de bit 2a à unidade de detecção de sinal 2e, a unidade de decodificação diferencial de coeficiente de previsão linear, e a unidade de geração de alta frequência 2g à unidade de conversão inversa de frequência 2n do dispositivo de decodificação de voz da modificação 2 são funções realizadas quando a CPU do dispositivo de decodificação de voz da modificação 2 executa o programa de computador armazenado na memória interna do dispositivo de decodificação de voz da modificação 2. Vários tipos de dados necessários para executar o programa de computador e vários tipos de dados gerados executando-se o programa de computador são todos armazenados na memória interna tal como a ROM e a RAM do dispositivo de decodificação de voz da modificação 2.

A unidade de decodificação diferencial de coeficiente de previsão linear obtém $a_{adj}(n, r)$ diferencialmente decodificado de acordo com a seguinte expressão (15), com uso de $a_L(n, r)$ obtido a partir da unidade de análise de previsão linear de baixa frequência 2d e $a_D(n, r)$ recebido a partir da unidade de separação de fluxo de bit 2a.

$$a_{adj}(n, r) = a_{dec}(n, r) + a_D(n, r), \quad 1 \leq n \leq N \quad \text{---(15)}$$

A unidade de decodificação diferencial de coeficiente de previsão linear transmite $a_{adj}(n, r)$ diferencialmente decodificado desta maneira para a unidade de filtro de previsão linear 2k. $a_D(n, r)$ pode ser um valor diferencial no domínio de coeficientes de previsão conforme ilustrado na expressão (14), mas pode ser um valor que toma uma diferença após converter um coeficiente de previsão à outra forma de expressão tal como LSP (Par de Espectros Lineares), ISP (Par de Espectros de Imitância), LSF (Frequência de Espectro Linear), ISF (Frequência de Espectro de Imitância), e coeficiente PARCOR. Neste caso, a decodificação diferencial também tem a mesma forma de expressão.

(Segunda modalidade)

A figura 6 é um diagrama que ilustra um dispositivo de codificação de voz 12 de acordo com a segunda modalidade. O dispositivo de codificação de voz 12 inclui fisicamente uma CPU, uma ROM, uma RAM, um dispositivo de comunicação, e similares, que não são ilustrados, e a CPU

5 controla integralmente o dispositivo de codificação de voz 12 carregando e executando um programa de computador predeterminado (tal como um programa de computador para realizar os processos ilustrados no fluxograma da figura 7) armazenado em uma memória interna do dispositivo de codificação de voz 12 tal como a ROM na RAM. O dispositivo de comunicação do

10 dispositivo de codificação de voz 12 recebe um sinal de voz a ser codificado de fora do dispositivo de codificação de voz 12, e emite um fluxo de bit multiplexado codificado para fora do dispositivo de codificação de voz 12.

O dispositivo de codificação de voz 12 funcionalmente inclui uma unidade de decimação de coeficiente de previsão linear 1j (meio de decimação de coeficiente de previsão), uma unidade de quantização de coeficiente de previsão 1k (meio de quantização de coeficiente de previsão), e a unidade de multiplexação de fluxo de bit 1g2 (meio de multiplexação de fluxo de bit), ao invés da unidade de cálculo de parâmetro de resistência de filtro 1f e da unidade de multiplexação de fluxo de bit 1g do dispositivo de codificação

20 de voz 11. A unidade de conversão de frequência 1a à unidade de análise de previsão linear 1e (meio de análise de previsão linear), a unidade de decimação de coeficiente de previsão linear 1j, a unidade de quantização de coeficiente de previsão 1k, e a unidade de multiplexação de fluxo de bit 1g2 do dispositivo de codificação de voz 12 ilustradas na figura 6 são funções

25 realizadas quando a CPU do dispositivo de codificação de voz 12 executa o programa de computador armazenado na memória interna do dispositivo de codificação de voz 12. A CPU do dispositivo de codificação de voz 12 sequencialmente executa processos (processos da Etapa Sa1 à Etapa Sa5, e processos da Etapa Sc1 à Etapa Sc3) ilustrados no fluxograma da figura 7,

30 executando-se o programa de computador (ou com uso da unidade de conversão de frequência 1a à unidade de análise de previsão linear 1e, a unidade de decimação de coeficiente de previsão linear 1j, a unidade de quantiza-

ção de coeficiente de previsão 1k, e a unidade de multiplexação de fluxo de bit 1g2 do dispositivo de codificação de voz 12 ilustrado na figura 6). Vários tipos de dados necessários para executar o programa de computador e vários tipos de dados gerados executando-se o programa de computador são
 5 todos armazenados na memória interna tal como a ROM e a RAM do dispositivo de codificação de voz 12.

A unidade de decimação de coeficiente de previsão linear 1j decima $a_H(n, r)$ obtido a partir da unidade de análise de previsão linear 1e na direção temporal, e transmite um valor para uma parte da abertura de tempo
 10 r_i e um valor que corresponde a r_i , em $a_H(n, r)$ para a unidade de quantização de coeficiente de previsão 1k (processo na Etapa Sc1). Nota-se que $0 \leq i < N_{ts}$, e N_{ts} é o número de aberturas de tempo às quais $a_H(n, r)$ é transmitido em um quadro. A decimação dos coeficientes de previsão linear pode ser realizada em um intervalo de tempo predeterminado, ou pode ser reali-
 15 zada em um intervalo de tempo irregular baseado nas características de $a_H(n, r)$. Por exemplo, um método é possível em que $G_H(r)$ de $a_H(n, r)$ é comparado em um quadro que tem um certo comprimento, e $a_H(n, r)$ é quantizado, se $G_H(r)$ exceder um certo valor. Se o intervalo de decimação do coeficiente de previsão linear for um intervalo predeterminado ao invés de usar as ca-
 20 racterísticas de $a_H(n, r)$, $a_H(n, r)$ não precisa ser calculado para a abertura de tempo para a qual a transmissão não é realizada.

A unidade de quantização de coeficiente de previsão 1k quantiza o coeficiente de previsão linear de alta frequência decimado $a_H(n, r_i)$ recebido a partir da unidade de decimação de coeficiente de previsão linear 1j e
 25 um índice r_i da abertura de tempo correspondente, e transmite para a unidade de multiplexação de fluxo de bit 1g2 (processo na Etapa Sc2). Como uma estrutura alternativa, ao invés de quantizar $a_H(n, r_i)$, um valor diferencial $a_D(n, r_i)$ dos coeficientes de previsão linear pode ser quantizado como o dispositivo de codificação de voz de acordo com a modificação 2 da primeira mo-
 30 dalidade.

A unidade de multiplexação de fluxo de bit 1g2 multiplexa o fluxo de bit codificado calculado pela unidade de codificação de codec central 1c,

as informações suplementares de SBR calculadas pela unidade de codificação de SBR 1d, e um índice $\{r_i\}$ de uma abertura de tempo que corresponde a $a_H(n, r_i)$ que é quantizado e recebido a partir da unidade de quantização de coeficiente de previsão 1k em um fluxo de bit, e emite o fluxo de bit multiplexado através do dispositivo de comunicação do dispositivo de codificação de voz 12 (processo na Etapa Sc3).

A figura 8 é um diagrama que ilustra o dispositivo de decodificação de voz 22 de acordo com a segunda modalidade. O dispositivo de decodificação de voz 22 inclui fisicamente uma CPU, uma ROM, uma RAM, um dispositivo de comunicação, e similares, que não são ilustrados, e a CPU controla integralmente o dispositivo de decodificação de voz 22 carregando e executando um programa de computador predeterminado (tal como um programa de computador para realizar os processos ilustrados no fluxograma da figura 9) armazenado em uma memória interna do dispositivo de decodificação de voz 22 tal como a ROM na RAM. O dispositivo de comunicação do dispositivo de decodificação de voz 22 recebe o fluxo de bit multiplexado codificado emitido a partir do dispositivo de codificação de voz 12, e emite um sinal de voz decodificado para fora do dispositivo de codificação de voz 12.

O dispositivo de decodificação de voz 22 funcionalmente inclui uma unidade de separação de fluxo de bit 2a1 (meio de separação de fluxo de bit), uma unidade de interpolação/extrapolação de coeficiente de previsão linear 2p (meio de interpolação/extrapolação de coeficiente de previsão linear), e uma unidade de filtro de previsão linear 2k1 (meio de transformação de envelope temporal) ao invés da unidade de separação de fluxo de bit 2a, da unidade de análise de previsão linear de baixa frequência 2d, da unidade de detecção de sinal 2e, da unidade de ajuste de resistência de filtro 2f, e da unidade de filtro de previsão linear 2k do dispositivo de decodificação de voz 21. A unidade de separação de fluxo de bit 2a1, a unidade de decodificação de codec central 2b, a unidade de conversão de frequência 2c, a unidade de geração de alta frequência 2g à unidade de ajuste de alta frequência 2j, a unidade de filtro de previsão linear 2k1, a unidade de adição de coeficiente

2m, a unidade de conversão inversa de frequência 2n, e a unidade de interpolação/extrapolação de coeficiente de previsão linear 2p do dispositivo de decodificação de voz 22 ilustrado na figura 8 são funções realizadas quando a CPU do dispositivo de codificação de voz 22 executa o programa de computador armazenado na memória interna do dispositivo de codificação de voz 22. A CPU do dispositivo de decodificação de voz 22 sequencialmente executa os processos (processos da Etapa Sb1 à Etapa Sd2, Etapa Sd1, da Etapa Sb5 à Etapa Sb8, Etapa Sd2, e da Etapa Sb10 à Etapa Sb11) ilustrados no fluxograma da figura 9, executando o programa de computador (ou com uso da unidade de separação de fluxo de bit 2a1, da unidade de decodificação de codec central 2b, da unidade de conversão de frequência 2c, da unidade de geração de alta frequência 2g à unidade de ajuste de alta frequência 2j, da unidade de filtro de previsão linear 2k1, da unidade de adição de coeficiente 2m, da unidade de conversão inversa de frequência 2n, e da unidade de interpolação/extrapolação de coeficiente de previsão linear 2p ilustrado na figura 8). Vários tipos de dados necessários para executar o programa de computador e vários tipos de dados gerados executando-se o programa de computador são todos armazenados na memória interna tal como a ROM e a RAM do dispositivo de decodificação de voz 22.

O dispositivo de decodificação de voz 22 inclui a unidade de separação de fluxo de bit 2a1, a unidade de interpolação/extrapolação de coeficiente de previsão linear 2p, e a unidade de filtro de previsão linear 2k1, ao invés da unidade de separação de fluxo de bit 2a, da unidade de análise de previsão linear de baixa frequência 2d, da unidade de detecção de sinal 2e, da unidade de ajuste de resistência de filtro 2f, e da unidade de filtro de previsão linear 2k do dispositivo de decodificação de voz 22.

A unidade de separação de fluxo de bit 2a1 separa o fluxo de bit multiplexado fornecido através do dispositivo de comunicação do dispositivo de decodificação de voz 22 no índice r_i da abertura de tempo que corresponde a $a_H(n, r_i)$ que é quantizado, as informações suplementares de SBR, e o fluxo de bit codificado.

A unidade de interpolação/extrapolação de coeficiente de previsão linear 2p recebe o índice r_i da abertura de tempo que corresponde a $a_H(n, r_i)$ que é quantizada a partir da unidade de separação de fluxo de bit 2a1, e obtém $a_H(n, r)$ que corresponde à abertura de tempo para a qual o coeficiente de previsão linear não é transmitido, por interpolação ou extrapolação (processos na Etapa Sd1). A unidade de interpolação/extrapolação de coeficiente de previsão linear 2p pode extrapolar o coeficiente de previsão linear, por exemplo, de acordo com a seguinte expressão (16).

$$a_H(n, r) = \delta^{|r-r_{i0}|} a_H(n, r_{i0}) \quad (1 \leq n \leq N) \quad \text{---(16)}$$

em que r_{i0} é o valor mais próximo de r na abertura de tempo $\{r_i\}$ para a qual o coeficiente de previsão linear é transmitido. δ é uma constante que satisfaz $0 < \delta < 1$.

A unidade de interpolação/extrapolação de coeficiente de previsão linear 2p pode interpolar o coeficiente de previsão linear, por exemplo, de acordo com a seguinte expressão (17), em que $r_{i0} < r < r_{i0+1}$ é satisfeito.

$$a_H(n, r) = \frac{r_{i0+1} - r}{r_{i0+1} - r_i} \cdot a_H(n, r_i) + \frac{r - r_{i0}}{r_{i0+1} - r_{i0}} \cdot a_H(n, r_{i0+1}) \quad (1 \leq n \leq N) \quad \text{---(17)}$$

A unidade de interpolação/extrapolação de coeficiente de previsão linear 2p pode converter o coeficiente de previsão linear em outras formas de expressão tal como LSP (Par de Espectros Lineares), ISP (Par de Espectros de Imitância), LSF (Frequência de Espectro Linear), ISF (Frequência de Espectro de Imitância), e coeficiente PARCOR, interpolar ou extrapolar este, e converter o valor obtido no coeficiente de previsão linear a ser usado. $a_H(n, r)$ que é interpolado ou extrapolado é transmitido para a unidade de filtro de previsão linear 2k1 e usado como um coeficiente de previsão linear para a filtragem de síntese de previsão linear, mas pode ser também usado como um coeficiente de previsão linear na unidade de filtro inverso de previsão linear 2i. Se $a_D(n, r_i)$ for multiplexado em um fluxo de bit ao invés de $a_H(n, r)$, a unidade de interpolação/extrapolação de coeficiente de previsão linear 2p realiza a decodificação diferencial similar àquela do dispositivo de decodificação de voz de acordo com a modificação 2 da pri-

meira modalidade, antes de realizar o processo de interpolação ou extrapolação descrito acima.

A unidade de filtro de previsão linear 2k1 realiza a filtragem de síntese de previsão linear na direção de frequência em $q_{adj}(n, r)$ emitido a partir da unidade de ajuste de alta frequência 2j, com uso de $a_H(n, r)$ que é interpolado ou extrapolado obtido a partir da unidade de interpolação/extrapolação de coeficiente de previsão linear 2p (processo na Etapa Sd2). Uma função de transferência da unidade de filtro de previsão linear 2k1 pode ser expressa como a seguinte expressão (18). A unidade de filtro de previsão linear 2k1 transforma o envelope temporal dos componentes de alta frequência gerados pela SBR realizando a filtragem de síntese de previsão linear, como a unidade de filtro de previsão linear 2k do dispositivo de decodificação de voz 21.

$$g(z) = \frac{1}{1 + \sum_{n=1}^N a_H(n, r) z^{-n}} \quad \text{---(18)}$$

(Terceira modalidade)

A figura 10 é um diagrama que ilustra um dispositivo de codificação de voz 13 de acordo com a terceira modalidade. O dispositivo de codificação de voz 13 inclui fisicamente uma CPU, uma ROM, uma RAM, um dispositivo de comunicação, e similares, que não são ilustrados, e a CPU controla integralmente o dispositivo de codificação de voz 13 carregando e executando um programa de computador predeterminado (tal como um programa de computador para realizar os processos ilustrados no fluxograma da figura 11) armazenado em uma memória interna do dispositivo de codificação de voz 13 tal como a ROM na RAM. O dispositivo de comunicação do dispositivo de codificação de voz 13 recebe um sinal de voz a ser codificado de fora do dispositivo de codificação de voz 13, e outputs um fluxo de bit multiplexado codificado para fora do dispositivo de codificação de voz 13.

O dispositivo de codificação de voz 13 funcionalmente inclui uma unidade de cálculo de envelope temporal 1m (meio de cálculo de informações suplementares de envelope temporal), uma unidade de cálculo de pa-

râmetro de formato de envelope 1n (meio de cálculo de informações suplementares de envelope temporal), e uma unidade de multiplexação de fluxo de bit 1g3 (meio de multiplexação de fluxo de bit), ao invés da unidade de análise de previsão linear 1e, da unidade de cálculo de parâmetro de resistência de filtro 1f, e da unidade de multiplexação de fluxo de bit 1g do dispositivo de codificação de voz 11. A unidade de conversão de frequência 1a à unidade de codificação de SBR 1d, a unidade de cálculo de envelope temporal 1m, a unidade de cálculo de parâmetro de formato de envelope 1n, e a unidade de multiplexação de fluxo de bit 1g3 do dispositivo de codificação de voz 13 ilustrado na figura 10 são funções realizadas quando a CPU do dispositivo de codificação de voz 13 executa o programa de computador armazenado na memória interna do dispositivo de codificação de voz 13. A CPU do dispositivo de codificação de voz 13 sequencialmente executa os processos (processos da Etapa Sa1 à Etapa Sa 4 e da Etapa Se1 à Etapa Se3) ilustrados no fluxograma da figura 11, executando o programa de computador (ou com uso da unidade de conversão de frequência 1a à unidade de codificação de SBR 1d, a unidade de cálculo de envelope temporal 1m, a unidade de cálculo de parâmetro de formato de envelope 1n, e a unidade de multiplexação de fluxo de bit 1g3 do dispositivo de codificação de voz 13 ilustrado na figura 10). Vários tipos de dados necessários para executar o programa de computador e vários tipos de dados gerados executando-se o programa de computador são todos armazenados na memória interna tal como a ROM e a RAM do dispositivo de codificação de voz 13.

A unidade de cálculo de envelope temporal 1m recebe $q(k, r)$, e, por exemplo, obtém informações de envelope temporal $e(r)$ dos componentes de alta frequência de um sinal, obtendo a força de cada abertura de tempo de $q(k, r)$ (processo na Etapa Se1). Neste caso, $e(r)$ é obtido de acordo com a seguinte expressão (19).

$$e(r) = \sqrt{\sum_{k=k_x}^{63} |q(k, r)|^2} \quad \text{---(19)}$$

A unidade de cálculo de parâmetro de formato de envelope 1n recebe $e(r)$ a partir da unidade de cálculo de envelope temporal 1m e recebe uma borda de tempo de envelope de SBR $\{b_i\}$ a partir da unidade de codificação de SBR 1d. Nota-se que $0 \leq i \leq N_e$, e N_e é o número de envelopes de SBR no quadro codificado. A unidade de cálculo de parâmetro de formato de envelope 1n obtém um parâmetro de formato de envelope $s(i)$ ($0 \leq i < N_e$) de cada um dos envelopes de SBR no quadro codificado de acordo com a seguinte expressão (20) (processo na Etapa Se2). O parâmetro de formato de envelope $s(i)$ corresponde às informações suplementares de envelope temporal, e é similar na terceira modalidade.

$$s(i) = \frac{1}{b_{i+1} - b_i - 1} \sum_{r=b_i}^{b_{i+1}-1} (\overline{e(i)} - e(r))^2$$

---(20)

Nota-se que:

$$\overline{e(i)} = \frac{\sum_{r=b_i}^{b_{i+1}-1} e(r)}{b_{i+1} - b_i} \quad \text{---(21)}$$

em que $s(i)$ na expressão acima é um parâmetro que indica a magnitude da variação de $e(r)$ no i° envelope de SBR que satisfaz $b_i \leq r < b_{i+1}$, e $e(r)$ toma um número maior conforme a variação do envelope temporal é aumentado. As expressões (20) e (21) descritas acima são exemplos de método para calcular $s(i)$, e, por exemplo, $s(i)$ pode ser também obtido com uso de, por exemplo, SMF (Medida de Planeza Espectral) de $e(r)$, uma razão do valor máximo para p valor mínimo, e similares. $s(i)$ é então quantizado e transmitido para a unidade de multiplexação de fluxo de bit 1g3.

A unidade de multiplexação de fluxo de bit 1g3 multiplexa o fluxo de bit codificado calculado pela unidade de codificação de codec central 1c, as informações suplementares de SBR calculadas pela unidade de codificação de SBR 1d, e $s(i)$ em um fluxo de bit, e emite o fluxo de bit multiplexado

através do dispositivo de comunicação do dispositivo de codificação de voz 13 (processo na Etapa Se3).

A figura 12 é um diagrama que ilustra um dispositivo de decodificação de voz 23 de acordo com a terceira modalidade. O dispositivo de decodificação de voz 23 inclui fisicamente uma CPU, uma ROM, uma RAM, um dispositivo de comunicação, e similares, que não são ilustrados, e a CPU controla integralmente o dispositivo de decodificação de voz 23 carregando e executando um programa de computador predeterminado (tal como um programa de computador para realizar os processos ilustrados no fluxograma da figura 13) armazenado em uma memória interna do dispositivo de decodificação de voz 23 tal como a ROM na RAM. O dispositivo de comunicação do dispositivo de decodificação de voz 23 recebe o fluxo de bit multiplexado codificado emitido a partir do dispositivo de codificação de voz 13, e emite um sinal de voz decodificado para fora do dispositivo de decodificação de voz 13.

O dispositivo de decodificação de voz 23 funcionalmente inclui uma unidade de separação de fluxo de bit 2a2 (meio de separação de fluxo de bit), uma unidade de cálculo de envelope temporal de baixa frequência 2r (meio de análise de envelope temporal de baixa frequência), uma unidade de ajuste de formato de envelope 2s (meio de ajuste de envelope temporal), uma unidade de cálculo de envelope temporal de alta frequência 2t, uma unidade de suavização de envelope temporal 2u, e uma unidade de transformação de envelope temporal 2v (meio de transformação de envelope temporal), ao invés da unidade de separação de fluxo de bit 2a, da unidade de análise de previsão linear de baixa frequência 2d, da unidade de detecção de sinal 2e, da unidade de ajuste de resistência de filtro 2f, a unidade de análise de previsão linear de alta frequência 2h, da unidade de filtro inverso de previsão linear 2i, e a unidade de filtro de previsão linear 2k do dispositivo de decodificação de voz 21. A unidade de separação de fluxo de bit 2a2, a unidade de decodificação de codec central 2b à unidade de conversão de frequência 2c, a unidade de geração de alta frequência 2g, a unidade de ajuste de alta frequência 2j, a unidade de adição de coeficiente 2m, a unidade

de conversão inversa de frequência 2n, e a unidade de cálculo de envelope temporal de baixa frequência 2r à unidade de transformação de envelope temporal 2v do dispositivo de decodificação de voz 23 ilustradas na figura 12 são funções realizadas quando a CPU do dispositivo de codificação de voz 23 executa o programa de computador armazenado na memória interna do dispositivo de codificação de voz 23. A CPU do dispositivo de decodificação de voz 23 sequencialmente executa os processos (processos da Etapa Sb1 à Etapa Sb2, da Etapa Sf1 à Etapa Sf2, Etapa Sb5, da Etapa Sf3 à Etapa Sf4, Etapa Sb8, Etapa Sf5, e da Etapa Sb10 à Etapa Sb11) ilustrados no fluxograma da figura 13, executando o programa de computador (ou com uso da unidade de separação de fluxo de bit 2a2, da unidade de decodificação de codec central 2b à unidade de conversão de frequência 2c, da unidade de geração de alta frequência 2g, da unidade de ajuste de alta frequência 2j, da unidade de adição de coeficiente 2m, da unidade de conversão inversa de frequência 2n, e da unidade de cálculo de envelope temporal de baixa frequência 2r à unidade de transformação de envelope temporal 2v do dispositivo de decodificação de voz 23 ilustradas na figura 12). Vários tipos de dados necessários para executar o programa de computador e vários tipos de dados gerados executando-se o programa de computador são todos armazenados na memória interna tal como a ROM e a RAM do dispositivo de decodificação de voz 23.

A unidade de separação de fluxo de bit 2a2 separa o fluxo de bit multiplexado fornecido através do dispositivo de comunicação do dispositivo de decodificação de voz 23 em $s(i)$, as informações suplementares de SBR e o fluxo de bit codificado. A unidade de cálculo de envelope temporal de baixa frequência 2r recebe $q_{dec}(k, r)$ que inclui os componentes de baixa frequência a partir da unidade de conversão de frequência 2c, e obtém $e(r)$ de acordo com a seguinte expressão (22) (processo na Etapa Sf1).

$$e(r) = \sqrt{\sum_{k=0}^{63} |q_{dec}(k, r)|^2} \quad \text{---(22)}$$

A unidade de ajuste de formato de envelope 2s ajusta $e(r)$ com uso de $s(i)$, e obtém as informações de envelope temporal ajustadas $e_{adj}(r)$

(processo na Etapa Sf2). $e(r)$ pode ser ajustado, por exemplo, de acordo com as seguintes expressões (23) a (25).

$$e_{adj}(r) = \overline{e(i)} + \sqrt{s(i) - v(i)} \cdot (e(r) - \overline{e(i)}) \quad (s(i) > v(i)) \quad \text{---(23)}$$

$$e_{adj}(r) = e(r) \quad \text{de outra forma}$$

Nota-se que:

$$\overline{e(i)} = \frac{\sum_{r=b_i}^{b_{i+1}-1} e(r)}{b_{i+1} - b_i} \quad \text{---(24)}$$

$$v(i) = \frac{1}{b_{i+1} - b_i - 1} \sum_{r=b_i}^{b_{i+1}-1} (\overline{e(i)} - e(r))^2 \quad \text{---(25)}$$

- As expressões (23) a (25) descritas acima são um exemplo de um método de ajuste, e o outro método de ajuste para o qual o formato de $e_{adj}(r)$ se torna similar ao formato ilustrado por (i) pode ser também usado.

A unidade de cálculo de envelope temporal de alta frequência 2t calcula um envelope temporal $e_{exp}(r)$ com uso de $q_{exp}(k, r)$ obtido a partir da unidade de geração de alta frequência 2g, de acordo com a seguinte expressão (26) (processo na Etapa Sf3).

$$e_{exp}(r) = \sqrt{\sum_{k=k_x}^{63} |q_{exp}(k, r)|^2} \quad \text{---(26)}$$

- A unidade de suavização de envelope temporal 2u suaviza o envelope temporal de $q_{exp}(k, r)$ obtido a partir da unidade de geração de alta frequência 2g de acordo com a seguinte expressão (27), e transmite o sinal obtido $q_{flat}(k, r)$ no domínio de QMF para a unidade de ajuste de alta frequência 2j (processo na Etapa Sf4).

$$q_{flat}(k, r) = \frac{q_{exp}(k, r)}{e_{exp}(r)} \quad (k_x \leq k \leq 63) \quad \text{---(27)}$$

- A suavização do envelope temporal pela unidade de suavização de envelope temporal 2u pode ser também omitida. Ao invés de calcular o envelope temporal dos componentes de alta frequência da saída a partir da unidade de geração de alta frequência 2g e suavizar o envelope temporal da mesma, o envelope temporal dos componentes de alta frequência de uma

saída a partir da unidade de ajuste de alta frequência 2j pode ser calculado, e o envelope temporal dos mesmos podem ser suavizados. O envelope temporal usado na unidade de suavização de envelope temporal 2u pode ser também $e_{adj}(r)$ obtido a partir da unidade de ajuste de formato de envelope 5 2s, ao invés de $e_{exp}(r)$ obtido a partir da unidade de cálculo de envelope temporal de alta frequência 2t.

A unidade de transformação de envelope temporal 2v transforma o $q_{adj}(k, r)$ obtido a partir da unidade de ajuste de alta frequência 2j com uso de $e_{adj}(r)$ obtido a partir da unidade de transformação de envelope temporal 10 2v, e obtém um sinal $q_{envadj}(k, r)$ no domínio de QMF em que o envelope temporal é transformado (processo na Etapa Sf5). A transformação é realizada de acordo com a seguinte expressão (28). $q_{envadj}(k, r)$ é transmitido para a unidade de adição de coeficiente 2m como um sinal no domínio de QMF que corresponde aos componentes de alta frequência.

$$q_{envadj}(k, r) = q_{adj}(k, r) \cdot e_{adj}(r) \quad (k_x \leq k \leq 63) \quad \text{---(28)}$$

15 (Quarta modalidade)

A figura 14 é um diagrama que ilustra um dispositivo de decodificação de voz 24 de acordo com a quarta modalidade. O dispositivo de decodificação de voz 24 inclui fisicamente uma CPU, uma ROM, uma RAM, um dispositivo de comunicação, e similares, que não são ilustrados, e a CPU 20 controla integralmente o dispositivo de decodificação de voz 24 carregando e executando a programa de computador predeterminado armazenado em uma memória interna do dispositivo de decodificação de voz 24 tal como a ROM na RAM. O dispositivo de comunicação do dispositivo de decodificação de voz 24 recebe o fluxo de bit multiplexado codificado emitido a partir do 25 dispositivo de codificação de voz 11 ou do dispositivo de codificação de voz 13, e emite um sinal de voz decodificado para fora do dispositivo de decodificação de voz 24.

O dispositivo de decodificação de voz 24 funcionalmente inclui a estrutura do dispositivo de decodificação de voz 21 (a unidade de decodificação de codec central 2b, a unidade de conversão de frequência 2c, a unidade de análise de previsão linear de baixa frequência 2d, a unidade de de- 30

tecção de sinal 2e, a unidade de ajuste de resistência de filtro 2f, a unidade de geração de alta frequência 2g, a unidade de análise de previsão linear de alta frequência 2h, a unidade de filtro inverso de previsão linear 2i, a unidade de ajuste de alta frequência 2j, a unidade de filtro de previsão linear 2k, a

5 unidade de adição de coeficiente 2m, e a unidade de conversão inversa de frequência 2n) e a estrutura do dispositivo de decodificação de voz 23 (a unidade de cálculo de envelope temporal de baixa frequência 2r, a unidade de ajuste de formato de envelope 2s, e a unidade de transformação de envelope temporal 2v). O dispositivo de decodificação de voz 24 também inclui

10 uma unidade de separação de fluxo de bit 2a3 (meio de separação de fluxo de bit) e uma unidade de conversão de informações suplementares 2w. A ordem da unidade de filtro de previsão linear 2k e da unidade de transformação de envelope temporal 2v pode ser oposta àquela ilustrada na figura 14. O dispositivo de decodificação de voz 24 preferencialmente recebe o fluxo

15 de bit codificado pelo dispositivo de codificação de voz 11 ou pelo dispositivo de codificação de voz 13. A estrutura do dispositivo de decodificação de voz 24 ilustrada na figura 14 é uma função realizada quando a CPU do dispositivo de decodificação de voz 24 executa o programa de computador armazenado na memória interna do dispositivo de decodificação de voz 24. Vários

20 tipos de dados necessários para executar o programa de computador e vários tipos de dados gerados executando-se o programa de computador são todos armazenados na memória interna tal como a ROM e a RAM do dispositivo de decodificação de voz 24.

A unidade de separação de fluxo de bit 2a3 separa o fluxo de bit

25 multiplexado fornecido através do dispositivo de comunicação do dispositivo de decodificação de voz 24 nas informações suplementares de envelope temporal, nas informações suplementares de SBR, e o fluxo de bit codificado. As informações suplementares de envelope temporal pode ser também $K(r)$ descrito na primeira modalidade ou $s(i)$ descrito na terceira modalidade.

30 As informações suplementares de envelope temporal podem ser também outro parâmetro $X(r)$ que não é nem $K(r)$ nem $s(i)$.

A unidade de conversão de informações suplementares 2w converte as informações suplementares de envelope temporal fornecidas para obter $K(r)$ e $s(i)$. Se as informações suplementares de envelope temporal for $K(r)$, a unidade de conversão de informações suplementares 2w converte

5 $K(r)$ em $s(i)$. A unidade de conversão de informações suplementares 2w pode também obter, por exemplo, um valor médio de $K(r)$ em uma seção de $b_i \leq r < b_{i+1}$

$$\overline{K(i)} \text{ ---(29)}$$

e converter o valor médio representado na expressão (29) em $s(i)$ com uso de uma tabela predeterminada. Se as informações suplementares de envelope temporal forem $s(i)$, a unidade de conversão de informações

10 suplementares 2w converte $s(i)$ em $K(r)$. A unidade de conversão de informações suplementares 2w pode também realizar a conversão convertendo $s(i)$ em $K(r)$, por exemplo, com uso de uma tabela predeterminada. Nota-se que i e r estão associados entre si assim como para satisfazer a relação de

15 $b_i \leq r < b_{i+1}$.

Se as informações suplementares de envelope temporal forem um parâmetro $X(r)$ que não é nem $s(i)$ nem $K(r)$, a unidade de conversão de informações suplementares 2w converte $X(r)$ em $K(r)$ e $s(i)$. É preferível que a unidade de conversão de informações suplementares 2w converta $X(r)$ em

20 $K(r)$ e $s(i)$, por exemplo, com uso de uma tabela predeterminada. É também preferível que a unidade de conversão de informações suplementares 2w transmita $X(r)$ para cada envelope de SBR como um valor representativo. As tabelas para converter $X(r)$ em $K(r)$ e $s(i)$ podem ser diferentes uma da outra.

(Modificação 3 da Primeira modalidade)

25 No dispositivo de decodificação de voz 21 da primeira modalidade, a unidade de filtro de previsão linear 2k do dispositivo de decodificação de voz 21 pode incluir um processo de controle de ganho automático. O processo de controle de ganho automático é um processo para ajustar a força do sinal no domínio de QMF emitido a partir da unidade de filtro de previsão

30 linear 2k para a força do sinal no domínio de QMF que é fornecido. Em geral,

um sinal $q_{syn,pow}(n, r)$ no domínio de QMF cujo ganho foi controlado é realizado pela seguinte expressão.

$$q_{syn,pow}(n, r) = q_{syn}(n, r) \cdot \sqrt{\frac{P_0(r)}{P_1(r)}} \quad \text{---(30)}$$

Aqui, $P_0(r)$ e $P_1(r)$ são expressos pela seguinte expressão (31) e a expressão (32).

$$P_0(r) = \sum_{n=k_x}^{63} |q_{adj}(n, r)|^2 \quad \text{---(31)}$$

$$P_1(r) = \sum_{n=k_x}^{63} |q_{syn}(n, r)|^2 \quad \text{---(32)}$$

- 5 Executando o processo de controle de ganho automático, a força dos componentes de alta frequência do sinal emitido a partir da unidade de filtro de previsão linear $2k$ é ajustada em um valor equivalente àquele antes da filtragem de previsão linear. Como resultado, no sinal emitido da unidade de filtro de previsão linear $2k$ em que o envelope temporal dos componentes
- 10 de alta frequência gerados com base em SBR é transformado, o efeito do ajuste da força do sinal de alta frequência realizado pela unidade de ajuste de alta frequência $2j$ pode ser mantido. O processo de controle de ganho automático pode ser também realizado individualmente em uma certa faixa de frequência do sinal no domínio de QMF. O processo realizado na faixa de
- 15 frequência individual pode ser realizado limitando-se n na expressão (30), a expressão (31), e a expressão (32) dentro de uma certa faixa de frequência. Por exemplo, a i^o faixa de frequência pode ser expressa como $F_i \leq n < F_{i+1}$ (neste caso, i é um índice que indica o número de uma certa faixa de frequência do sinal no domínio de QMF). F_i indica o limite da faixa de frequên-
- 20 cia, e é preferível que F_i seja uma tabela de limite de frequência de um fator de escala de envelope definido em SBR em "MPEG4 AAC". A tabela de limite de frequência é definida pela unidade de geração de alta frequência $2g$ com base na definição de SBR em "MPEG4 AAC". Realizando-se o processo de controle de ganho automático, a força do sinal emitido a partir da uni-
- 25 dade de filtro de previsão linear $2k$ em uma certa faixa de frequência dos

componentes de alta frequência é ajustada em um valor equivalente àquele antes da filtragem de previsão linear. Como resultado, o efeito para ajustar a força do sinal de alta frequência realizado pela unidade de ajuste de alta frequência 2j no sinal emitido a partir da unidade de filtro de previsão linear 2k em que o envelope temporal dos componentes de alta frequência gerados com base em SBR é transformado, é mantido por unidade de faixa de frequência. As alterações feitas na presente modificação 3 da primeira modalidade podem também ser feitas na unidade de filtro de previsão linear 2k da quarta modalidade.

10 [Modificação 1 da Terceira modalidade]

A unidade de cálculo de parâmetro de formato de envelope 1n no dispositivo de codificação de voz 13 da terceira modalidade pode ser também realizada pelos seguintes processos. A unidade de cálculo de parâmetro de formato de envelope 1n obtém um parâmetro de formato de envelope s(i) ($0 \leq i < N_e$) de acordo com a seguinte expressão (33) para cada envelope de SBR no quadro codificado.

$$s(i) = 1 - \min\left(\frac{e(r)}{e(i)}\right) \quad \text{---(33)}$$

Nota-se que:

$$\frac{e(r)}{e(i)} \quad \text{---(34)}$$

é um valor médio de e(r) no envelope de SBR, e o método de cálculo é baseado na expressão (21). Nota-se que o envelope de SBR indica a faixa de tempo que satisfaz $b_i \leq r < b_{i+1}$. $\{b_i\}$ é a borda de tempo de envelope de SBR incluída nas informações suplementares de SBR como as informações, e é o limite da faixa de tempo para a qual o fator de escala de envelope de SBR que representa a energia de sinal média em uma certa faixa de tempo e uma certa faixa de frequência é dada. min (·) representa o valor mínimo dentro da faixa de $b_i \leq r < b_{i+1}$. Desta forma, neste caso, o parâmetro de formato de envelope s(i) é um parâmetro para indicar uma razão do valor mínimo para o valor médio no envelope de SBR das informações de envelope temporal ajustadas. A unidade de ajuste de formato de envelope 2s no dispositivo de decodificação de voz 23 da terceira modalidade pode ser tam-

bém realizada pelo seguinte processo. A unidade de ajuste de formato de envelope 2s ajusta $e(r)$ com uso de $s(i)$ para obter as informações de envelope temporal ajustadas $e_{adj}(r)$. O método de ajuste é baseado na seguinte expressão (35) ou expressão (36).

$$e_{adj}(r) = \overline{e(i)} \left(1 + s(i) \frac{(e(r) - \overline{e(i)})}{\overline{e(i)} - \min(e(r))} \right) \quad \text{---(35)}$$

$$e_{adj}(r) = \overline{e(i)} \left(1 + s(i) \frac{(e(r) - \overline{e(i)})}{\overline{e(i)}} \right) \quad \text{---(36)}$$

- 5 A expressão 35 ajusta o formato do envelope de forma que a razão do valor mínimo para o valor médio no envelope de envelope de SBR das informações de envelope temporal ajustadas $e_{adj}(r)$ se torne equivalente ao valor do parâmetro de formato de envelope $s(i)$. As alterações feitas na modificação 1 da terceira modalidade descritas acima podem ser também
- 10 feitas na quarta modalidade.

[Modificação 2 da Terceira modalidade]

- A unidade de transformação de envelope temporal 2v pode também usar a seguinte expressão ao invés da expressão (28). Conforme indicado na expressão (37), $e_{adj, \text{scaled}}(r)$ é obtido controlando-se o ganho das
- 15 informações de envelope temporal ajustadas $e_{adj}(r)$, de forma que a força de $q_{adj}(k, r)$ e $q_{envadj}(k, r)$ sejam equivalentes entre si no envelope de SBR. Conforme indicado na expressão (38), na presente modificação 2 da terceira modalidade, $q_{envadj}(k, r)$ é obtido multiplicando-se $e_{adj, \text{scaled}}(r)$ ao invés de $e_{adj}(r)$ pelo sinal $q_{adj}(k, r)$ no domínio de QMF. Da mesma forma, a unidade
- 20 de transformação de envelope temporal 2v pode transformar o envelope temporal do sinal $q_{adj}(k, r)$ no domínio de QMF, de forma que a força do sinal no envelope de SBR se torne equivalente antes e após a transformação do envelope temporal. Nota-se que o envelope de SBR indica a faixa de tempo que satisfaz $b_i \leq r < b_{i+1}$. $\{b_i\}$ é a borda de tempo de envelope de SBR incluída
- 25 nas informações suplementares de SBR como as informações, e é o limite da faixa de tempo para a qual o fator de escala de envelope de SBR que representa a energia de sinal média de uma certa faixa de tempo e uma certa faixa de frequência é dada. A terminologia "envelope de SBR" nas modali-

dades da presente invenção corresponde à terminologia "segmento de tempo de envelope de SBR" em "MPEG4 AAC" definido em "ISO/IEC 14496-3", e o "envelope de SBR" tem os mesmos conteúdos que o "segmento de tempo de envelope de SBR" por todas as modalidades.

$$e_{adj,scaled}(r) = e_{adj}(r) \cdot \sqrt{\frac{\sum_{k=k_x}^{63} \sum_{r=b_i}^{b_{i+1}-1} |q_{adj}(k,r)|^2}{\sum_{k=k_x}^{63} \sum_{r=b_i}^{b_{i+1}-1} |q_{adj}(k,r) \cdot e_{adj}(r)|^2}} \quad ---(37)$$

$$(k_x \leq k \leq 63, b_i \leq r < b_{i+1})$$

$$q_{envadj}(k,r) = q_{adj}(k,r) \cdot e_{adj,scaled}(r) \quad ---(38)$$

$$(k_x \leq k \leq 63, b_i \leq r < b_{i+1})$$

5

As alterações feitas na presente modificação 2 a terceira modalidade descritas acima podem ser também feitas na quarta modalidade.

(Modificação 3 da Terceira modalidade)

A expressão (19) pode ser também a seguinte expressão (39).

$$e(r) = \sqrt{\frac{(b_{i+1} - b_i) \sum_{k=0}^{63} |q(k,r)|^2}{\sum_{r=b_i}^{b_{i+1}-1} \sum_{k=0}^{63} |q(k,r)|^2}} \quad ---(39)$$

A expressão (22) por $k=k_x$... $k=k_x$, a seguinte expressão (40).

$$e(r) = \sqrt{\frac{(b_{i+1} - b_i) \sum_{k=0}^{63} |q_{dec}(k,r)|^2}{\sum_{r=b_i}^{b_{i+1}-1} \sum_{k=0}^{63} |q_{dec}(k,r)|^2}} \quad ---(40)$$

10

A expressão (26) pode ser também a seguinte expressão (41).

$$e_{exp}(r) = \sqrt{\frac{(b_{i+1} - b_i) \sum_{k=k_x}^{63} |q_{exp}(k,r)|^2}{\sum_{r=b_i}^{b_{i+1}-1} \sum_{k=k_x}^{63} |q_{exp}(k,r)|^2}} \quad ---(41)$$

Quando a expressão (39) e a expressão (40) são usadas, as in-

5 formações de envelope temporal $e(r)$ são informações em que a força de
 cada amostra de sub-banda de QMF é normalizada pela força média no en-
 velope de SBR, e a raiz quadrada é extraída. No entanto, a amostra de sub-
 banda de QMF é um vetor de sinal que corresponde ao índice de tempo " r "
 10 que é o mesmo no domínio de QMF sinal, e é uma sub-amostra no domínio
 de QMF. Em todas as modalidades da presente invenção, a terminologia
 "abertura de tempo" tem os mesmos conteúdos que a "amostra de sub-
 banda de QMF". Neste caso, as informações de envelope temporal $e(r)$ são
 um coeficiente de ganho que deve ser multiplicado por cada amostra de sub-
 10 banda de QMF, e o mesmo se aplica às informações de envelope temporal
 ajustadas $e_{adj}(r)$.

(Modificação 1 da Quarta modalidade)

Um dispositivo de decodificação de voz 24a (não ilustrado) de
 uma modificação 1 da quarta modalidade inclui fisicamente uma CPU, uma
 15 ROM, uma RAM, um dispositivo de comunicação, e similares, que não são
 ilustrados, e a CPU controla integralmente o dispositivo de decodificação de
 voz 24a carregando e executando um programa de computador predetermi-
 nado armazenado em uma memória interna do dispositivo de decodificação
 de voz 24a tal como a ROM na RAM. O dispositivo de comunicação do dis-
 20 positivo de decodificação de voz 24a recebe o fluxo de bit multiplexado codi-
 ficado emitido a partir do dispositivo de codificação de voz 11 ou do disposi-
 tivo de codificação de voz 13, e emite um sinal de voz decodificado para fora
 do dispositivo de decodificação de voz 24a. O dispositivo de decodificação
 de voz 24a funcionalmente inclui uma unidade de separação de fluxo de bit
 25 2a4 (não ilustrado) ao invés da unidade de separação de fluxo de bit 2a3 do
 dispositivo de decodificação de voz 24, e também inclui a unidade de gera-
 ção de informações suplementares de envelope temporal 2y (não ilustrado),
 ao invés da unidade de conversão de informações suplementares 2w. A uni-
 dade de separação de fluxo de bit 2a4 separa o fluxo de bit multiplexado nas
 30 informações de SBR e o fluxo de bit codificado. A unidade de geração de
 informações suplementares de envelope temporal 2y gera informações su-

plementares de envelope temporal com base nas informações incluídas no fluxo de bit codificado e as informações suplementares de SBR.

Para gerar as informações suplementares de envelope temporal em um certo envelope de SBR, por exemplo, a largura de tempo ($b_{i+1}-b_i$) do envelope de SBR, uma classe de quadro, um parâmetro de resistência do filtro inverso, a base de ruído, a amplitude da força de alta frequência, uma razão da força de alta frequência para a força de baixa frequência, um coeficiente de auto-correlação ou um ganho de previsão obtido como resultado de realizar a análise de previsão linear na direção de frequência em um sinal de baixa frequência representado no domínio de QMF, e similares podem ser usados. As informações suplementares de envelope temporal podem ser geradas determinando $K(r)$ ou $s(i)$ com base em um ou uma pluralidade de valores dos parâmetros. Por exemplo, as informações suplementares de envelope temporal podem ser geradas determinando $K(r)$ ou $s(i)$ com base em ($b_{i+1}-b_i$) de forma que $K(r)$ ou $s(i)$ seja reduzido conforme a largura de tempo ($b_{i+1}-b_i$) do envelope de SBR é aumentada, ou $K(r)$ ou $s(i)$ seja aumentado conforme a largura de tempo ($b_{i+1}-b_i$) do envelope de SBR é aumentada. As alterações similares podem ser também feitas na primeira modalidade e na terceira modalidade.

20 (Modificação 2 da Quarta modalidade)

Um dispositivo de decodificação de voz 24b (vide a figura 15) de uma modificação 2 da quarta modalidade inclui fisicamente uma CPU, uma ROM, uma RAM, um dispositivo de comunicação, e similares, que não são ilustrados, e a CPU controla integralmente o dispositivo de decodificação de voz 24b carregando e executando um programa de computador predeterminado armazenado em uma memória interna do dispositivo de decodificação de voz 24b tal como a ROM na RAM. O dispositivo de comunicação do dispositivo de decodificação de voz 24b recebe o fluxo de bit multiplexado codificado emitido a partir do dispositivo de codificação de voz 11 ou do dispositivo de codificação de voz 13, e emite um sinal de voz decodificado para fora do dispositivo de decodificação de voz 24b. O dispositivo de decodificação de voz 24b, conforme ilustrado na figura 15, inclui uma unidade de ajuste de

alta frequência primária 2j1 e uma unidade de ajuste de alta frequência secundária 2j2 ao invés da unidade de ajuste de alta frequência 2j.

Aqui, a unidade de ajuste de alta frequência primária 2j1 ajusta um sinal no domínio de QMF a banda de alta frequência realizando o ajuste, e a superposição de ruído, descrito na etapa de "ajuste de HF" em SBR em "MPEG4 AAC". Neste momento, o sinal emitido da unidade de ajuste de alta frequência primária 2j1 corresponde a um sinal W_2 na descrição em "ferramenta de SBR" em "ISO/IEC 14496-3:2005", cláusulas 4.6.18.7.6 de "Sinais de HF de montagem". A unidade de filtro de previsão linear 2k (ou a unidade de filtro de previsão linear 2k1) e a unidade de transformação de envelope temporal 2v transformam o envelope temporal do sinal emitido a partir da unidade de ajuste de alta frequência primária. A unidade de ajuste de alta frequência secundária 2j2 realiza um processo de adição de onda de seno na etapa de "ajuste de HF" em SBR em "MPEG4 AAC". O processo da unidade de ajuste de alta frequência secundária corresponde a um processo para substituir o sinal W_2 por um sinal emitido da unidade de transformação de envelope temporal 2v, em um processo para gerar um sinal Y a partir do sinal W_2 na descrição em "ferramenta de SBR" em "ISO/IEC 14496-3:2005", cláusulas 4.6.18.7.6 de "Sinais de HF de montagem".

Na descrição acima, somente o processo para adicionar ondas de seno é realizado pela unidade de ajuste de alta frequência secundária 2j2. No entanto, qualquer um dos processos na etapa de "ajuste de HF" pode ser realizado pela unidade de ajuste de alta frequência secundária 2j2. Modificações similares podem ser também feitas na primeira modalidade, na segunda modalidade e na terceira modalidade. Neste momento, a unidade de filtro de previsão linear (unidades de filtro de previsão linear 2k e 2k1) é incluída na primeira modalidade e na segunda modalidade, mas a unidade de transformação de envelope temporal não é incluída. Da mesma forma, um sinal emitido a partir da unidade de ajuste de alta frequência primária 2j1 é processado pela unidade de filtro de previsão linear, e então um sinal emitido a partir da unidade de filtro de previsão linear é processado pela unidade de ajuste de alta frequência secundária 2j2.

Na terceira modalidade, a unidade de transformação de envelope temporal 2v está incluída, mas a unidade de filtro de previsão linear não está incluída. Da mesma forma, um sinal emitido a partir da unidade de ajuste de alta frequência primária 2j1 é processado pela unidade de transformação de envelope temporal 2v, e então um sinal emitido a partir da unidade de transformação de envelope temporal 2v é processado pela unidade de ajuste de alta frequência secundária.

No dispositivo de decodificação de voz (dispositivo de decodificação de voz 24, 24a, ou 24b) da quarta modalidade, a ordem de processamento da unidade de filtro de previsão linear 2k e da unidade de transformação de envelope temporal 2v pode ser revertida. Em outras palavras, um sinal emitido a partir da unidade de ajuste de alta frequência 2j ou da unidade de ajuste de alta frequência primária 2j1 pode ser processado primeiro pela unidade de transformação de envelope temporal 2v, e então um sinal emitido a partir da unidade de transformação de envelope temporal 2v pode ser processado pela unidade de filtro de previsão linear 2k.

Em adição, somente se as informações suplementares de envelope temporal incluir informações de controle binário para indicar se o processo é realizado pela unidade de filtro de previsão linear 2k ou pela unidade de transformação de envelope temporal 2v, e as informações de controle indicam para realizar o processo pela unidade de filtro de previsão linear 2k ou pela unidade de transformação de envelope temporal 2v, as informações suplementares de envelope temporal podem empregar uma forma que inclui ao menos um dentre o parâmetro de resistência de filtro $K(r)$, o parâmetro de formato de envelope $s(i)$, ou $X(r)$ que é um parâmetro para determinar ambos $K(r)$ e $s(i)$ como informações.

(Modificação 3 da Quarta modalidade)

Um dispositivo de decodificação de voz 24c (vide a figura 16) de uma modificação 3 da quarta modalidade inclui fisicamente uma CPU, uma ROM, uma RAM, um dispositivo de comunicação, e similares, que não são ilustrados, e a CPU controla integralmente o dispositivo de decodificação de voz 24c carregando e executando um programa de computador predetermi-

nado (tal como um programa de computador para realizar os processos ilustrados no fluxograma da figura 17) armazenado em uma memória interna do dispositivo de decodificação de voz 24c tal como a ROM na RAM. O dispositivo de comunicação do dispositivo de decodificação de voz 24c recebe o

5 fluxo de bit multiplexado codificado e emite um sinal de voz decodificado para fora do dispositivo de decodificação de voz 24c. Conforme ilustrado na figura 16, o dispositivo de decodificação de voz 24c inclui uma unidade de ajuste de alta frequência primária 2j3 e uma unidade de ajuste de alta frequência secundária 2j4 ao invés da unidade de ajuste de alta frequência 2j, e

10 também inclui unidades de ajuste de componente de sinal individual 2z1, 2z2, e 2z3 ao invés da unidade de filtro de previsão linear 2k e da unidade de transformação de envelope temporal 2v (as unidades de ajuste de componente de sinal individual correspondem ao meio de transformação de envelope temporal).

15 A unidade de ajuste de alta frequência primária 2j3 emite um sinal no domínio de QMF da banda de alta frequência como um componente de sinal copiado. A unidade de ajuste de alta frequência primária 2j3 pode emitir um sinal em que ao menos um dentre a filtragem inversa de previsão linear na direção temporal e o ajuste de ganho (ajuste de características de

20 frequência) é realizado no sinal no domínio de QMF da banda de alta frequência, com uso das informações suplementares de SBR recebidas a partir da unidade de separação de fluxo de bit 2a3, como um componente de sinal copiado. A unidade de ajuste de alta frequência primária 2j3 também gera um componente de sinal de ruído e um componente de sinal de onda de se-

25 no com uso das informações suplementares de SBR fornecidas a partir da unidade de separação de fluxo de bit 2a3, e emite cada um dentre o componente de sinal copiado, o componente de sinal de ruído, e o componente de sinal de onda de seno em uma forma separada (processo na Etapa Sg1). O componente de sinal de ruído e o componente de sinal de onda de seno po-

30 dem não ser gerados, dependendo dos conteúdos das informações suplementares de SBR.

As unidades de ajuste de componente de sinal individual 2z1,

2z2, e 2z3 realizam o processamento em cada um na pluralidade de componentes de sinal incluídos na saída da unidade de ajuste de alta frequência primária (processo na Etapa Sg2). O processo com as unidades de ajuste de componente de sinal individual 2z1, 2z2, e 2z3 pode ser filtragem de síntese

5 de previsão linear na direção de frequência obtida a partir da unidade de ajuste de resistência de filtro 2f com uso do coeficiente de previsão linear, similar àquele da unidade de filtro de previsão linear 2k (processo 1). O processo com as unidades de ajuste de componente de sinal individual 2z1, 2z2, e 2z3 pode ser também um processo para multiplicar um coeficiente de

10 ganho por cada amostra de sub-banda de QMF com uso do envelope temporal obtido a partir da unidade de ajuste de formato de envelope 2s, similar àquele da unidade de transformação de envelope temporal 2v (processo 2). O processo com as unidades de ajuste de componente de sinal individual 2z1, 2z2, e 2z3 podem ser também um processo para realizar a filtragem de

15 síntese de previsão linear na direção de frequência no sinal de entrada com uso do coeficiente de previsão linear obtido a partir da unidade de ajuste de resistência de filtro 2f similar àquele da unidade de filtro de previsão linear 2k, e então multiplicar um coeficiente de ganho por cada amostra de sub-banda de QMF com uso do envelope temporal obtido a partir da unidade de

20 ajuste de formato de envelope 2s, similar àquele da unidade de transformação de envelope temporal 2v (processo 3). O processo com as unidades de ajuste de componente de sinal individual 2z1, 2z2, e 2z3 pode ser também um processo para multiplicar um coeficiente de ganho por cada amostra de sub-banda de QMF com relação ao sinal de entrada com uso do envelope

25 temporal obtido a partir da unidade de ajuste de formato de envelope 2s, similar àquele da unidade de transformação de envelope temporal 2v, e então realizar a filtragem de síntese de previsão linear na direção de frequência no sinal emitido com uso do coeficiente de previsão linear obtido a partir da unidade de ajuste de resistência de filtro 2f, similar àquele da unidade de

30 filtro de previsão linear 2k (processo 4). As unidades de ajuste de componente de sinal individual 2z1, 2z2, e 2z3 podem não realizar o processo de transformação de envelope temporal no sinal de entrada, mas pode emitir o

5 sinal de entrada como é (processo 5). O processo com as unidades de ajuste de componente de sinal individual 2z1, 2z2, e 2z3 pode incluir qualquer processo para transformar o envelope temporal do sinal de entrada com uso de um método outro que os processos 1 a 5 (processo 6). O processo com as unidades de ajuste de componente de sinal individual 2z1, 2z2, e 2z3 pode ser também um processo em que uma pluralidade de processos entre os processos 1 a 6 são combinados em uma ordem arbitrária (processo 7).

10 Os processos com as unidades de ajuste de componente de sinal individual 2z1, 2z2, e 2z3 podem ser os mesmos, mas as unidades de ajuste de componente de sinal individual 2z1, 2z2, e 2z3 podem transformar o envelope temporal de cada componente da pluralidade de componentes de sinal incluídos na saída da unidade de ajuste de alta frequência primária por diferentes métodos. Por exemplo, diferentes processos podem ser realizados no sinal copiado, no sinal de ruído, e no sinal de onda de seno, de tal
15 maneira que a unidade de ajuste de componente de sinal individual 2z1 realize o processo 2 no sinal copiado fornecido, a unidade de ajuste de componente de sinal individual 2z2 realize o processo 3 no componente de sinal de ruído fornecido, e a unidade de ajuste de componente de sinal individual 2z3 realize o processo 5 no sinal de onda de seno fornecido. Neste momento, a
20 unidade de ajuste de resistência de filtro 2f e a unidade de ajuste de formato de envelope 2s podem transmitir o mesmo coeficiente de previsão linear e o envelope temporal para as unidades de ajuste de componente de sinal individual 2z1, 2z2, e 2z3, mas podem também transmitir diferentes coeficientes de previsão linear e os envelopes temporais. É também possível transmitir o
25 mesmo coeficiente de previsão linear e o envelope temporal para ao menos duas as unidades de ajuste de componente de sinal individual 2z1, 2z2, e 2z3. Devido ao fato de que ao menos um das unidades de ajuste de componente de sinal individual 2z1, 2z2, e 2z3 podem não realizar o processo de transformação de envelope temporal, mas emitirem o sinal de entrada como
30 este é (processo 5), as unidades de ajuste de componente de sinal individual 2z1, 2z2, e 2z3 realizam o processo de envelope temporal em ao menos um componente da pluralidade de componentes de sinal emitidos a partir da

unidade de ajuste de alta frequência primária 2j3 como um todo (se todas as unidades de ajuste de componente de sinal individual 2z1, 2z2, e 2z3 realizarem o processo 5, o processo de transformação de envelope temporal não é realizado em qualquer um dos componentes de sinal, e os efeitos da presente invenção não são exibidos).

Os processos realizados por cada uma das unidades de ajuste de componente de sinal individual 2z1, 2z2, e 2z3 podem ser fixados em um dentre o processo 1 ao processo 7, mas podem ser dinamicamente determinado para realizar um dentre o processo 1 ao processo 7 com base nas informações de controle recebidas de fora do dispositivo de decodificação de voz 24c. Neste momento, é preferível que as informações de controle estejam incluídas no fluxo de bit multiplexado. As informações de controle podem ser uma instrução para realizar qualquer um dentre o processo 1 ao processo 7 em um segmento de tempo de envelope de SBR específico, o quadro codificado, ou na outra faixa de tempo, ou podem ser uma instrução para realizar qualquer um dentre o processo 1 ao processo 7 sem especificar a faixa de tempo de controle.

A unidade de ajuste de alta frequência secundária 2j4 adiciona os componentes de sinal processados emitidos a partir das unidades de ajuste de componente de sinal individual 2z1, 2z2, e 2z3, e emite o resultado para a unidade de adição de coeficiente (processo na Etapa Sg3). A unidade de ajuste de alta frequência secundária 2j4 pode realizar ao menos um dentre a filtragem inversa de previsão linear na direção temporal e ajuste de ganho (ajuste de característica de frequência) no componente de sinal copiado, com uso das informações suplementares de SBR recebidas a partir da unidade de separação de fluxo de bit 2a3.

As unidades de ajuste de componente de sinal individual 2z1, 2z2, e 2z3 podem operar em cooperação entre si, e gerar um sinal omitido em um estágio intermediário adicionando ao menos dois componentes de sinal em que qualquer um dos processos 1 a 7 é realizado, e ainda realizar qualquer um dos processos 1 a 7 no sinal adicionado. Neste momento, a unidade de ajuste de alta frequência secundária 2j4 adiciona o sinal emitido

no estágio intermediário e um componente de sinal que não foi ainda adicionado ao sinal emitido no estágio intermediário, e emite o resultado para a unidade de adição de coeficiente. Mais especificamente, é preferível gerar um sinal emitido no estágio intermediário realizando o processo 5 no componente de sinal copiado, aplicando o processo 1 no componente de ruído, adicionando dois componentes de sinal, e ainda aplicando o processo 2 no sinal adicionado. Neste momento, a unidade de ajuste de alta frequência secundária 2j4 adiciona o componente de sinal de onda de seno ao sinal emitido no estágio intermediário, e emite o resultado para a unidade de adição de coeficiente.

A unidade de ajuste de alta frequência primária 2j3 pode emitir qualquer um de uma pluralidade de componentes de sinal em uma forma separada entre si em adição aos três componentes de sinal do componente de sinal copiado, do componente de sinal de ruído, e do componente de sinal de onda de seno. Neste caso, o componente de sinal pode ser obtido adicionando-se ao menos dois dentre o componente de sinal copiado, o componente de sinal de ruído, e o componente de sinal de onda de seno. O componente de sinal pode ser também um sinal obtido dividindo-se a banda de um dentre o componente de sinal copiado, o componente de sinal de ruído, e o sinal de onda de seno. O número de componentes de sinal pode ser outro que não três, e neste caso, o número das unidades de ajuste de componente de sinal individual pode ser outro que não três.

O sinal de alta frequência gerado por SBR é formado de três elementos do componente de sinal copiado obtido copiando da banda de baixa frequência para a banda de alta frequência, o sinal de ruído, e o sinal de onda de seno. Devido ao fato de que o sinal copiado, o sinal de ruído, e o sinal de onda de seno têm os envelopes temporais diferentes entre si, se o envelope temporal de cada um dos componentes de sinal for transformado com uso dos métodos diferentes como as unidades de ajuste de componente de sinal individual da presente modificação, é possível melhorar ainda mais a qualidade subjetiva do sinal decodificado comparado com as outras modalidades da presente invenção. Em particular, devido ao fato de que o

sinal de ruído de forma geral tem um envelope temporal suave, e o sinal co-
 piado tem um envelope temporal próximo àquele do sinal na banda de baixa
 frequência, os envelopes temporais do sinal copiado e o sinal de ruído po-
 dem ser independentemente controlados, tratando-os separadamente e apli-
 cando diferentes processos a estes. Da mesma forma, é efetivo em melhorar
 a qualidade subjetiva do sinal decodificado. Mais especificamente, é preferí-
 vel realizar um processo para transformar o envelope temporal no sinal de
 ruído (processo 3 ou processo 4), realizar um processo diferente daquele
 para o sinal de ruído no sinal copiado (processo 1 ou processo 2), e realizar
 o processo 5 no sinal de onda de seno (em outras palavras, o processo de
 transformação de envelope temporal não é realizado). É também preferível
 realizar um processo de transformação (processo 3 ou processo 4) do enve-
 lope temporal no sinal de ruído, e realizar o processo 5 no sinal copiado e o
 sinal de onda de seno (em outras palavras, o processo de transformação de
 envelope temporal não é realizado).

(Modificação 4 da Primeira modalidade)

Um dispositivo de codificação de voz 11b (figura 44) de uma
 modificação 4 da primeira modalidade inclui fisicamente uma CPU, uma
 ROM, uma RAM, um dispositivo de comunicação, e similares, que não são
 ilustrados, e a CPU controla integralmente o dispositivo de codificação de
 voz 11b carregando e executando um programa de computador predetermi-
 nado armazenado em uma memória interna do dispositivo de codificação de
 voz 11b tal como a ROM na RAM. O dispositivo de comunicação do disposi-
 tivo de codificação de voz 11b recebe um sinal de voz a ser codificado de
 fora do dispositivo de codificação de voz 11b, e emite um fluxo de bit multi-
 plexado codificado para fora do dispositivo de codificação de voz 11b. The
 dispositivo de codificação de voz 11b includes a unidade de análise de previ-
 são linear 1e1 ao invés da unidade de análise de previsão linear 1e do dis-
 positivo de codificação de voz 11, e ainda inclui uma unidade de seleção de
 abertura de tempo 1p.

A unidade de seleção de abertura de tempo 1p recebe um sinal
 no domínio de QMF a partir da unidade de conversão de frequência 1a e

seleciona uma abertura de tempo na qual a análise de previsão linear pela unidade de análise de previsão linear 1e1 é realizada. A unidade de análise de previsão linear 1e1 realiza a análise de previsão linear no sinal de domínio de QMF na abertura de tempo selecionada como a unidade de análise

5 de previsão linear 1e, com base no resultado de seleção transmitido a partir da unidade de seleção de abertura de tempo 1p, para obter ao menos um dentre o coeficiente de previsão linear de alta frequência e o coeficiente de previsão linear de baixa frequência. A unidade de cálculo de parâmetro de resistência de filtro 1f calcula um parâmetro de resistência de filtro com uso

10 de um coeficiente de previsão linear da abertura de tempo selecionado pela unidade de seleção de abertura de tempo 1p, obtido pela unidade de análise de previsão linear 1e1. Para selecionar uma abertura de tempo pela unidade de seleção de abertura de tempo 1p, por exemplo, ao menos um dos métodos de seleção com uso da força de sinal do sinal de domínio de QMF dos

15 componentes de alta frequência, similares àqueles de uma unidade de seleção de abertura de tempo 3a em um dispositivo de decodificação 21a da presente modificação, que será descrito a diante, pode ser usado. Neste momento, é preferível que o sinal de domínio de QMF dos componentes de alta frequência na unidade de seleção de abertura de tempo 1p seja um

20 componente de frequência codificado pela unidade de codificação de SBR 1d, entre os sinais no domínio de QMF recebido a partir da unidade de conversão de frequência 1a. O método de seleção de abertura de tempo pode ser ao menos um dos métodos descritos acima, pode incluir ao menos um método diferente aqueles descritos acima, ou pode ser a combinação dos

25 mesmos.

Um dispositivo de decodificação de voz 21a (vide a figura 18) da modificação 4 da primeira modalidade inclui fisicamente uma CPU, uma ROM, uma RAM, um dispositivo de comunicação, e similares, que não são ilustrados, e a CPU controla integralmente o dispositivo de decodificação de

30 voz 21a carregando e executando um programa de computador predeterminado (tal como um programa de computador para realizar os processos ilustrados no fluxograma da figura 19) armazenado em uma memória interna do

dispositivo de decodificação de voz 21a tal como a ROM na RAM. O dispositivo de comunicação do dispositivo de decodificação de voz 21a recebe o fluxo de bit multiplexado codificado e emite um sinal de voz decodificado para fora do dispositivo de decodificação de voz 21a. O dispositivo de decodificação de voz 21a, conforme ilustrado na figura 18, inclui uma unidade de análise de previsão linear de baixa frequência 2d1, uma unidade de detecção de sinal 2e1, uma unidade de análise de previsão linear de alta frequência 2h1, uma unidade de filtro inverso de previsão linear 2i1, e uma unidade de filtro de previsão linear 2k3 ao invés da unidade de análise de previsão linear de baixa frequência 2d, da unidade de detecção de sinal 2e, da unidade de análise de previsão linear de alta frequência 2h, da unidade de filtro inverso de previsão linear 2i, e a unidade de filtro de previsão linear 2k do dispositivo de decodificação de voz 21, e ainda inclui a unidade de seleção de abertura de tempo 3a.

A unidade de seleção de abertura de tempo 3a determina se a filtragem de síntese de previsão linear na unidade de filtro de previsão linear 2k deve ser realizada no sinal $q_{exp}(k, r)$ no domínio de QMF dos componentes de alta frequência da abertura de tempo r gerados pela unidade de geração de alta frequência 2g, e seleciona uma abertura de tempo em que a filtragem de síntese de previsão linear é realizada (processo na Etapa Sh1). A unidade de seleção de abertura de tempo 3a notifica, do resultado de seleção da abertura de tempo, a unidade de análise de previsão linear de baixa frequência 2d1, a unidade de detecção de sinal 2e1, a unidade de análise de previsão linear de alta frequência 2h1, a unidade de filtro inverso de previsão linear 2i1, e a unidade de filtro de previsão linear 2k3. A unidade de análise de previsão linear de baixa frequência 2d1 realiza a análise de previsão linear no sinal de domínio de QMF na abertura de tempo selecionada $r1$, da mesma maneira que a unidade de análise de previsão linear de baixa frequência 2d, com base no resultado de seleção transmitido a partir da unidade de seleção de abertura de tempo 3a, para obter um coeficiente de previsão linear de baixa frequência (processo na Etapa Sh2). A unidade de detecção de sinal 2e1 detecta a variação temporal no sinal de domínio de QMF na

abertura de tempo selecionada, como a unidade de detecção de sinal 2e, com base no resultado de seleção transmitido a partir da unidade de seleção de abertura de tempo 3a, e emite um resultado de detecção T (r1).

A unidade de ajuste de resistência de filtro 2f realiza o ajuste de
 5 resistência de filtro no coeficiente de previsão linear de baixa frequência da abertura de tempo selecionada pela unidade de seleção de abertura de tempo 3a obtida pela unidade de análise de previsão linear de baixa frequência 2d1, para obter um coeficiente de previsão linear ajustado $a_{dec}(n, r1)$. A unidade de análise de previsão linear de alta frequência 2h1 realiza a análise
 10 de previsão linear na direção de frequência no sinal de domínio de QMF dos componentes de alta frequência gerados pela unidade de geração de alta frequência 2g para a abertura de tempo selecionada r1, com base no resultado de seleção transmitido a partir da unidade de seleção de abertura de tempo 3a, como a unidade de análise de previsão linear de alta frequência
 15 2h, para obter um coeficiente de previsão linear de alta frequência $a_{exp}(n, r1)$ (processo na Etapa Sh3). A unidade de filtro inverso de previsão linear 2i1 realiza a filtragem inversa de previsão linear, em que $a_{exp}(n, r1)$ é um coeficiente, na direção de frequência no sinal $q_{exp}(k, r)$ no domínio de QMF dos componentes de alta frequência da abertura de tempo selecionada r1, como
 20 a unidade de filtro inverso de previsão linear 2i, com base no resultado de seleção transmitido a partir da unidade de seleção de abertura de tempo 3a (processo na Etapa Sh4).

A unidade de filtro de previsão linear 2k3 realiza a filtragem de síntese de previsão linear na direção de frequência em um sinal $q_{adj}(k, r1)$ no
 25 domínio de QMF dos componentes de alta frequência emitidos a partir da unidade de ajuste de alta frequência 2j na abertura de tempo selecionada r1 com uso de $a_{adj}(n, r1)$ obtido a partir da unidade de ajuste de resistência de filtro 2f, como a unidade de filtro de previsão linear 2k, com base no resultado de seleção transmitido a partir da unidade de seleção de abertura de tempo 3a (processo na Etapa Sh5). As alterações feitas na unidade de filtro
 30 de previsão linear 2k descrita na modificação 3 podem ser também feitas na unidade de filtro de previsão linear 2k3. Para selecionar uma abertura de

tempo em que a filtragem de síntese de previsão linear é realizada, por exemplo, a unidade de seleção de abertura de tempo 3a pode selecionar ao menos uma abertura de tempo r em que a força de sinal do sinal de domínio de QMF $q_{\text{exp}}(k, r)$ dos componentes de alta frequência é maior que um valor
 5 predeterminado $P_{\text{exp,Th}}$. É preferível calcular a força de sinal de $q_{\text{exp}}(k, r)$ de acordo com a seguinte expressão.

$$P_{\text{exp}}(r) = \sum_{k=k_x}^{k_x+M-1} |q_{\text{exp}}(k, r)|^2 \quad \text{---(42)}$$

em que M é um valor que representa uma faixa de frequência mais alta que uma frequência limite inferior k_x dos componentes de alta frequência gerados pela unidade de geração de alta frequência $2g$, e a faixa de frequência dos
 10 componentes de alta frequência gerados pela unidade de geração de alta frequência $2g$ pode ser representada como $k_x \leq k < k_x + M$. O valor predeterminado $P_{\text{exp,Th}}$ pode ser também um valor médio de uma largura de tempo predeterminada $P_{\text{exp}}(r)$ incluindo a abertura de tempo r . A largura de tempo predeterminada pode ser também o envelope de SBR.

15 A seleção pode ser também feita de forma a incluir uma abertura de tempo em que a força de sinal do sinal de domínio de QMF dos componentes de alta frequência alcance seu pico. A força de sinal de pico pode ser calculada, por exemplo, com uso de um valor de média em movimento:

$$P_{\text{exp,MA}}(r) \quad \text{---(43)}$$

da força de sinal, e a força de sinal de pico pode ser a força de sinal no domínio de QMF dos componentes de alta frequência da abertura de tempo r
 20 em que o resultado de:

$$P_{\text{exp,MA}}(r+1) - P_{\text{exp,MA}}(r) \quad \text{---(44)}$$

altera do valor positivo para o valor negativo. O valor de média em movimento da força de sinal,

$$P_{\text{exp,MA}}(r) \quad \text{---(45)}$$

por exemplo, pode ser calculada pela seguinte expressão.

$$P_{\text{exp,MA}}(r) = \frac{1}{c} \sum_{r'=r-\frac{c}{2}}^{r+\frac{c}{2}-1} P_{\text{exp}}(r') \quad \text{---(46)}$$

em que c é um valor predeterminado para definir uma faixa para calcular o valor médio. A força de sinal de pico pode ser calculada pelo método descrito acima, ou pode ser calculada por um método diferente.

Ao menos uma abertura de tempo pode ser selecionada a partir das aberturas de tempo incluídas em uma largura de tempo t durante a qual a força de sinal do sinal de domínio de QMF dos componentes de alta frequência é alterada de um estado estável com uma pequena variação para um estado transiente com uma grande variação, e que são menores que um valor predeterminado t_{th} . Ao menos uma abertura de tempo pode ser também selecionada de aberturas de tempo incluídas em uma largura de tempo t durante a qual a força de sinal do sinal de domínio de QMF dos componentes de alta frequência é alterada de um estado transiente com uma grande variação para um estado estável com uma pequena variação, e que são maiores que o valor predeterminado t_{th} . A abertura de tempo r em que $|P_{exp}(r+1)-P_{exp}(r)|$ é menor que um valor predeterminado (ou igual ou menor que um valor predeterminado) pode ser o estado estável, e a abertura de tempo r em que $|P_{exp}(r+1)-P_{exp}(r)|$ é igual ou maior que um valor predeterminado (ou maior que um valor predeterminado) pode ser o estado transiente. A abertura de tempo r em que $|P_{exp,MA}(r+1)-P_{exp,MA}(r)|$ é menor que um valor predeterminado (ou igual ou menor que um valor predeterminado) pode ser o estado estável, e a abertura de tempo r em que $|P_{exp,MA}(r+1)-P_{exp,MA}(r)|$ é igual ou maior que um valor predeterminado (ou maior que um valor predeterminado) pode ser o estado transiente. O estado transiente e o estado estável podem ser definidos com uso do método descrito acima, ou podem ser definidos com uso de diferentes métodos. O método de seleção de abertura de tempo pode ser ao menos um dos métodos descritos acima, pode incluir ao menos um método diferente daqueles descritos acima, ou pode ser a combinação dos mesmos.

(Modificação 5 da Primeira modalidade)

Um dispositivo de codificação de voz 11c (figura 45) de uma modificação 5 da primeira modalidade inclui fisicamente uma CPU, uma ROM, uma RAM, um dispositivo de comunicação, e similares, que não são ilustra-

dos, e a CPU controla integralmente o dispositivo de codificação de voz 11c carregando e executando um programa de computador predeterminado armazenado em uma memória interna do dispositivo de codificação de voz 11c tal como a ROM na RAM. O dispositivo de comunicação do dispositivo de

5 codificação de voz 11c recebe um sinal de voz a ser codificado de fora do dispositivo de codificação de voz 11c, e emite um fluxo de bit multiplexado codificado para fora do dispositivo de codificação de voz 11c. O dispositivo de codificação de voz 11c inclui uma unidade de seleção de abertura de tempo 1p1 e uma unidade de multiplexação de fluxo de bit 1g4, ao invés da

10 unidade de seleção de abertura de tempo 1p e da unidade de multiplexação de fluxo de bit 1g do dispositivo de codificação de voz 11b da modificação 4.

A unidade de seleção de abertura de tempo 1p1 seleciona uma abertura de tempo como a unidade de seleção de abertura de tempo 1p descrita na modificação 4 da primeira modalidade, e transmite informações

15 de seleção de tempo de abertura para a unidade de multiplexação de fluxo de bit 1g4. A unidade de multiplexação de fluxo de bit 1g4 multiplexa o fluxo de bit codificado calculado pela unidade de codificação de codec central 1c, as informações suplementares de SBR calculadas pela unidade de codificação de SBR 1d, e o parâmetro de resistência de filtro calculado pela unidade

20 de cálculo de parâmetro de resistência de filtro 1f como a unidade de multiplexação de fluxo de bit 1g, também multiplexa as informações de seleção de tempo de abertura recebidas a partir da unidade de seleção de abertura de tempo 1p1, e emite o fluxo de bit multiplexado através do dispositivo de comunicação do dispositivo de codificação de voz 11c. As informações de

25 seleção de tempo de abertura são informações de seleção de tempo de abertura recebidas por uma unidade de seleção de abertura de tempo 3a1 em um dispositivo de decodificação de voz 21b, que será descrito adiante, e, por exemplo, um índice r1 de uma abertura de tempo a ser selecionada pode ser incluído. As informações de seleção de tempo de abertura podem ser tam-

30 bém um método de seleção de parâmetro usado na abertura de tempo da unidade de seleção de abertura de tempo 3a1. O dispositivo de decodificação de voz 21b (vide a figura 20) da modificação 5 da primeira modalidade

inclui fisicamente uma CPU, uma ROM, uma RAM, um dispositivo de comunicação, e similares, que não são ilustrados, e a CPU controla integralmente o dispositivo de decodificação de voz 21b carregando e executando um programa de computador predeterminado (tal como um programa de computador para realizar os processos ilustrados no fluxograma a figura 21) armazenado em uma memória interna do dispositivo de decodificação de voz 21b tal como a ROM na RAM. O dispositivo de comunicação do dispositivo de decodificação de voz 21b recebe o fluxo de bit multiplexado codificado e emite um sinal de voz decodificado para fora do dispositivo de decodificação de voz 21b.

O dispositivo de decodificação de voz 21b, conforme ilustrado na figura 20, inclui uma unidade de separação de fluxo de bit 2a5 e a unidade de seleção de abertura de tempo 3a1 ao invés da unidade de separação de fluxo de bit 2a e da unidade de seleção de abertura de tempo 3a do dispositivo de decodificação de voz 21a da modificação 4, e as informações de seleção de tempo de abertura são fornecidas para a unidade de seleção de abertura de tempo 3a1. A unidade de separação de fluxo de bit 2a5 separa o fluxo de bit multiplexado no parâmetro de resistência de filtro, as informações suplementares de SBR, e o fluxo de bit codificado como a unidade de separação de fluxo de bit 2a, e ainda separa as informações de seleção de tempo de abertura. A unidade de seleção de abertura de tempo 3a1 seleciona uma abertura de tempo com base nas informações de seleção de tempo de abertura transmitidas a partir da unidade de separação de fluxo de bit 2a5 (processo na Etapa Si1). As informações de seleção de tempo de abertura são informações usadas para selecionar uma abertura de tempo, e, por exemplo, pode incluir o índice r1 da abertura de tempo a ser selecionada. As informações de seleção de tempo de abertura podem ser também um parâmetro, por exemplo, usado no método de seleção de abertura de tempo na modificação 4. Neste caso, embora não ilustrado, o sinal de domínio de QMF dos componentes de alta frequência gerados pela unidade de geração de alta frequência 2g pode ser fornecido para a unidade de seleção de abertura de tempo 3a1, em adição às informações de seleção de tempo de abertura.

O parâmetro pode ser também um valor predeterminado (tal como $P_{exp,Th}$ e t_{Th}) usado para selecionar a abertura de tempo.

(Modificação 6 da Primeira modalidade)

Um dispositivo de codificação de voz 11d (não ilustrado) de uma
 5 modificação 6 da primeira modalidade inclui fisicamente uma CPU, uma ROM, uma RAM, um dispositivo de comunicação, e similares, que não são ilustrados, e a CPU controla integralmente o dispositivo de codificação de voz 11d carregando e executando um programa de computador predeterminado armazenado em uma memória interna do dispositivo de codificação de
 10 voz 11d tal como a ROM na RAM. O dispositivo de comunicação do dispositivo de codificação de voz 11d recebe um sinal de voz a ser codificado de fora do dispositivo de codificação de voz 11d, e emite um fluxo de bit multiplexado codificado para fora do dispositivo de codificação de voz 11d. O dispositivo de codificação de voz 11d inclui uma unidade de cálculo de força de
 15 período curto 1i1, que não é ilustrada, ao invés da unidade de cálculo de força de período curto 1i do dispositivo de codificação de voz 11a da modificação 1, e ainda inclui uma unidade de seleção de abertura de tempo 1p2.

A unidade de seleção de abertura de tempo 1p2 recebe um sinal no domínio de QMF a partir da unidade de conversão de frequência 1a, e
 20 seleciona uma abertura de tempo que corresponde à seção de tempo em que o processo de cálculo de força de período curto é realizado pela unidade de cálculo de força de período curto 1i. A unidade de cálculo de força de período curto 1i1 calcula a força de período curto de uma seção de tempo que corresponde à abertura de tempo selecionada com base no resultado de
 25 seleção transmitido a partir da unidade de seleção de abertura de tempo 1p2, como a unidade de cálculo de força de período curto 1i do dispositivo de codificação de voz 11a da modificação 1.

(Modificação 7 da Primeira modalidade)

Um dispositivo de codificação de voz 11e (não ilustrado) de uma
 30 modificação 7 a primeira modalidade inclui fisicamente uma CPU, uma ROM, uma RAM, u a dispositivo de comunicação, e similares, que não são ilustrados, e a CPU controla integralmente o dispositivo de codificação de voz 11e

carregando e executando um programa de computador predeterminado armazenado em uma memória interna do dispositivo de codificação de voz 11e tal como a ROM na RAM. O dispositivo de comunicação do dispositivo de codificação de voz 11e recebe um sinal de voz a ser codificado de fora do dispositivo de codificação de voz 11e, e emite um fluxo de bit multiplexado codificado para fora do dispositivo de codificação de voz 11e. O dispositivo de codificação de voz 11e inclui uma unidade de seleção de abertura de tempo 1p3, que não é ilustrada, ao invés da unidade de seleção de abertura de tempo 1p2 do dispositivo de codificação de voz 11d da modificação 6. O dispositivo de codificação de voz 11e também inclui uma unidade de multiplexação de fluxo de bit que ainda recebe uma saída da unidade de seleção de abertura de tempo 1p3, ao invés da unidade de multiplexação de fluxo de bit 1g1. A unidade de seleção de abertura de tempo 1p3 seleciona uma abertura de tempo como a unidade de seleção de abertura de tempo 1p2 descrita na modificação 6 da primeira modalidade, e transmite as informações de seleção de tempo de abertura para a unidade de multiplexação de fluxo de bit.

(Modificação 8 da Primeira modalidade)

Um dispositivo de codificação de voz (não ilustrado) de uma modificação 8 da primeira modalidade inclui fisicamente uma CPU, uma ROM, uma RAM, um dispositivo de comunicação, e similares, que não são ilustrados, e a CPU controla integralmente o dispositivo de codificação de voz da modificação 8 carregando e executando um programa de computador predeterminado armazenado em uma memória interna do dispositivo de codificação de voz da modificação 8 tal como a ROM na RAM. O dispositivo de comunicação do dispositivo de codificação de voz da modificação 8 recebe um sinal de voz a ser codificado de fora do dispositivo de codificação de voz, e emite um fluxo de bit multiplexado codificado para fora do dispositivo de codificação de voz. O dispositivo de codificação de voz da modificação 8 ainda inclui a unidade de seleção de abertura de tempo 1p em adição àquelas do dispositivo de codificação de voz descrito na modificação 2.

Um dispositivo de decodificação de voz (não ilustrado) da modificação 8 da primeira modalidade inclui fisicamente uma CPU, uma ROM, uma RAM, um dispositivo de comunicação, e similares, que não são ilustrados, e a CPU controla integralmente o dispositivo de decodificação de voz da modificação 8 carregando e executando um programa de computador predeterminado armazenado em uma memória interna do dispositivo de decodificação de voz da modificação 8 tal como a ROM na RAM. O dispositivo de comunicação do dispositivo de decodificação de voz da modificação 8 recebe o fluxo de bit multiplexado codificado, e emite um sinal de voz decodificado para fora do dispositivo de decodificação de voz. O dispositivo de decodificação de voz da modificação 8 ainda inclui a unidade de análise de previsão linear de baixa frequência 2d1, a unidade de detecção de sinal 2e1, a unidade de análise de previsão linear de alta frequência 2h1, a unidade de filtro inverso de previsão linear 2i1, e a unidade de filtro de previsão linear 2k3, ao invés da unidade de análise de previsão linear de baixa frequência 2d, da unidade de detecção de sinal 2e, da unidade de análise de previsão linear de alta frequência 2h, da unidade de filtro inverso de previsão linear 2i, e da unidade de filtro de previsão linear 2k do dispositivo de decodificação de voz descrito na modificação 2, e ainda inclui a unidade de seleção de abertura de tempo 3a.

(Modificação 9 da Primeira modalidade)

Um dispositivo de codificação de voz (não ilustrado) de uma modificação 9 da primeira modalidade inclui fisicamente uma CPU, uma ROM, uma RAM, um dispositivo de comunicação, e similares, que não são ilustrados, e a CPU controla integralmente o dispositivo de codificação de voz da modificação 9 carregando e executando um programa de computador predeterminado armazenado em uma memória interna do dispositivo de codificação de voz da modificação 9 tal como a ROM na RAM. O dispositivo de comunicação do dispositivo de codificação de voz da modificação 9 recebe um sinal de voz a ser codificado de fora do dispositivo de codificação de voz, e emite um fluxo de bit multiplexado codificado para fora do dispositivo de codificação de voz. O dispositivo de codificação de voz da modificação 9 inclui

a unidade de seleção de abertura de tempo 1p1 ao invés da unidade de seleção de abertura de tempo 1p do dispositivo de codificação de voz descrito na modificação 8. O dispositivo de codificação de voz da modificação 9 ainda inclui uma unidade de multiplexação de fluxo de bit que recebe uma saída da unidade de seleção de abertura de tempo 1p1 em adição para entrada fornecida para a unidade de multiplexação de fluxo de bit descrita na modificação 8, ao invés da unidade de multiplexação de fluxo de bit descrita na modificação 8.

Um dispositivo de decodificação de voz (não ilustrado) da modificação 9 da primeira modalidade inclui fisicamente uma CPU, uma ROM, uma RAM, um dispositivo de comunicação, e similares, que não são ilustrados, e a CPU controla integralmente o dispositivo de decodificação de voz da modificação 9 carregando e executando um programa de computador predeterminado armazenado em uma memória interna do dispositivo de decodificação de voz da modificação 9 tal como a ROM na RAM. O dispositivo de comunicação do dispositivo de decodificação de voz da modificação 9 recebe o fluxo de bit multiplexado codificado, e emite um sinal de voz decodificado para fora do dispositivo de decodificação de voz. O dispositivo de decodificação de voz da modificação 9 inclui a unidade de seleção de abertura de tempo 3a1 ao invés da unidade de seleção de abertura de tempo 3a do dispositivo de decodificação de voz descrito na modificação 8. O dispositivo de decodificação de voz da modificação 9 ainda inclui uma unidade de separação de fluxo de bit que separa $a_D(n, r)$ descrito na modificação 2 ao invés do parâmetro de resistência de filtro da unidade de separação de fluxo de bit 2a5, ao invés da unidade de separação de fluxo de bit 2a.

(Modificação 1 da Segunda modalidade)

Um dispositivo de codificação de voz 12a (figura 46) de uma modificação 1 da segunda modalidade inclui fisicamente uma CPU, uma ROM, uma RAM, um dispositivo de comunicação, e similares, que não são ilustrados, e a CPU controla integralmente o dispositivo de codificação de voz 12a carregando e executando um programa de computador predeterminado armazenado em uma memória interna do dispositivo de codificação de

voz 12a tal como a ROM na RAM. O dispositivo de comunicação do dispositivo de codificação de voz 12a recebe um sinal de voz a ser codificado de fora do dispositivo de codificação de voz 12a, e emite um fluxo de bit multiplexado codificado para fora do dispositivo de codificação de voz 12a.. O

5 dispositivo de codificação de voz 12a inclui a unidade de análise de previsão linear 1e1 ao invés da unidade de análise de previsão linear 1e do dispositivo de codificação de voz 12, e ainda inclui a unidade de seleção de abertura de tempo 1p.

Um dispositivo de decodificação de voz 22a (vide a figura 22) da

10 modificação 1 da segunda modalidade inclui fisicamente uma CPU, uma ROM, uma RAM, um dispositivo de comunicação, e similares, que não são ilustrados, e a CPU controla integralmente o dispositivo de decodificação de voz 22a carregando e executando um programa de computador predeterminado (tal como um programa de computador para realizar os processos ilus-

15 trados no fluxograma da figura 23) armazenado em uma memória interna do dispositivo de decodificação de voz 22a tal como a ROM na RAM. O dispositivo de comunicação do dispositivo de decodificação de voz 22a recebe o fluxo de bit multiplexado codificado, e emite um sinal de voz decodificado para fora do dispositivo de decodificação de voz 22a. O dispositivo de deco-

20 dificação de voz 22a, conforme ilustrado na figura 22, inclui a unidade de análise de previsão linear de alta frequência 2h1, a unidade de filtro inverso de previsão linear 2i1, uma unidade de filtro de previsão linear 2k2, e uma unidade de interpolação/extrapolação de previsão linear 2p1, ao invés da unidade de análise de previsão linear de alta frequência 2h, a unidade de

25 filtro inverso de previsão linear 2i, a unidade de filtro de previsão linear 2k1, e a unidade de interpolação/extrapolação de previsão linear 2p do dispositivo de decodificação de voz 22 da segunda modalidade, e ainda inclui a unidade de seleção de abertura de tempo 3a.

A unidade de seleção de abertura de tempo 3a notifica, do resul-

30 tado de seleção da abertura de tempo, a unidade de análise de previsão linear de alta frequência 2h1a a unidade de filtro inverso de previsão linear 2i1, a unidade de filtro de previsão linear 2k2, e a unidade de interpola-

ção/extrapolação de coeficiente de previsão linear 2p1. A unidade de interpolação/extrapolação de coeficiente de previsão linear 2p1 obtém $a_H(n, r)$ que corresponde à abertura de tempo $r1$ que é a abertura de tempo selecionada e para a qual um coeficiente de previsão linear não é transmitido por interpolação ou extrapolação, como a unidade de interpolação/extrapolação de coeficiente de previsão linear 2p, com base no resultado de seleção transmitido a partir da unidade de seleção de abertura de tempo 3a (processo na Etapa Sj1). A unidade de filtro de previsão linear 2k2 realiza a filtragem de síntese de previsão linear na direção de frequência em $q_{adj}(n, r1)$ emitido a partir da unidade de ajuste de alta frequência 2j para a abertura de tempo selecionada $r1$ com uso de $a_H(n, r1)$ que é interpolado ou extrapolado e obtido a partir da unidade de interpolação/extrapolação de coeficiente de previsão linear 2p1, como a unidade de filtro de previsão linear 2k1 (processo na Etapa Sj2), com base no resultado de seleção transmitido a partir da unidade de seleção de abertura de tempo 3a. As alterações feitas na unidade de filtro de previsão linear 2k descrita na modificação 3 da primeira modalidade podem ser também feitas na unidade de filtro de previsão linear 2k2.

(Modificação 2 da Segunda modalidade)

Um dispositivo de codificação de voz 12b (figura 47) de uma modificação 2 da segunda modalidade inclui fisicamente uma CPU, uma ROM, uma RAM, um dispositivo de comunicação, e similares, que não são ilustrados, e a CPU controla integralmente o dispositivo de codificação de voz 11b carregando e executando um programa de computador predeterminado armazenado em uma memória interna do dispositivo de codificação de voz 12b tal como a ROM na RAM. O dispositivo de comunicação do dispositivo de codificação de voz 12b recebe um sinal de voz a ser codificado de fora do dispositivo de codificação de voz 12b, e emite um fluxo de bit multiplexado codificado para fora do dispositivo de codificação de voz 12b. O dispositivo de codificação de voz 12b inclui a unidade de seleção de abertura de tempo 1p1 e uma unidade de multiplexação de fluxo de bit 1g5 ao invés da unidade de seleção de abertura de tempo 1p e da unidade de multiplexação de fluxo de bit 1g2 do dispositivo de codificação de voz 12a da modifica-

ção 1. A unidade de multiplexação de fluxo de bit 1g5 multiplexa o fluxo de bit codificado calculado pela unidade de codificação de codec central 1c, as informações suplementares de SBR calculadas pela unidade de codificação de SBR 1d, e um índice da abertura de tempo que corresponde ao coeficiente de previsão linear quantizado recebido a partir da unidade de quantização de coeficiente de previsão 1k como a unidade de multiplexação de fluxo de bit 1g2, ainda multiplexa as informações de seleção de tempo de abertura recebidas a partir da unidade de seleção de abertura de tempo 1p1, e emite o fluxo de bit multiplexado através do dispositivo de comunicação do dispositivo de codificação de voz 12b.

Um dispositivo de decodificação de voz 22b (vide a figura 24) da modificação 2 da segunda modalidade inclui fisicamente uma CPU, uma ROM, uma RAM, um dispositivo de comunicação, e similares, que não são ilustrados, e a CPU controla integralmente o dispositivo de decodificação de voz 22b carregando e executando um programa de computador predeterminado (tal como um programa de computador para realizar os processos ilustrados no fluxograma da figura 25) armazenado em uma memória interna do dispositivo de decodificação de voz 22b tal como a ROM na RAM. O dispositivo de comunicação do dispositivo de decodificação de voz 22b recebe o fluxo de bit multiplexado codificado, e emite um sinal de voz decodificado para fora do dispositivo de decodificação de voz 22b. O dispositivo de decodificação de voz 22b, conforme ilustrado na figura 24, inclui uma unidade de separação de fluxo de bit 2a6 e a unidade de seleção de abertura de tempo 3a1 ao invés da unidade de separação de fluxo de bit 2a1 e da unidade de seleção de abertura de tempo 3a do dispositivo de decodificação de voz 22a descrito na modificação 1, e as informações de seleção de tempo de abertura são fornecidas para a unidade de seleção de abertura de tempo 3a1. A unidade de separação de fluxo de bit 2a6 separa o fluxo de bit multiplexado em $a_H(n, r_i)$ que é quantizado, o índice r_i da abertura de tempo correspondente, as informações suplementares de SBR, e o fluxo de bit codificado como a unidade de separação de fluxo de bit 2a1, e ainda separa as informações de seleção de tempo de abertura.

(Modificação 4 da Terceira modalidade)

$$\overline{e(i)} \quad \text{---(47)}$$

descrita na modificação 1 da terceira modalidade pode ser um valor médio de $e(r)$ no envelope de SBR, ou pode ser um valor definido de alguma outra maneira.

5 (Modificação 5 da Terceira modalidade)

Conforme descrito na modificação 3 da terceira modalidade, é preferível que a unidade de ajuste de formato de envelope 2s controle $e_{adj}(r)$ com uso de um valor predeterminado $e_{adj,Th}(r)$, considerando-se que o envelope temporal ajustado $e_{adj}(r)$ é um coeficiente de ganho multiplicado pela amostra de sub-banda de QMF, por exemplo, como a expressão (28) e as expressões (37) e (38).

$$e_{adj}(r) \geq e_{adj,Th} \quad \text{---(48)}$$

(Quarta modalidade)

Um dispositivo de codificação de voz 14 (FIGURA 48) da quarta modalidade inclui fisicamente uma CPU, uma ROM, uma RAM, um dispositivo de comunicação, e similares, que não são ilustrados, e a CPU controla integralmente o dispositivo de codificação de voz 14 carregando e executando um programa de computador predeterminado armazenado em uma memória interna do dispositivo de codificação de voz 14 tal como a ROM na RAM. O dispositivo de comunicação do dispositivo de codificação de voz 14 recebe um sinal de voz a ser codificado de fora do dispositivo de codificação de voz 14, e emite um fluxo de bit multiplexado codificado para fora do dispositivo de codificação de voz 14. O dispositivo de codificação de voz 14 inclui uma unidade de multiplexação de fluxo de bit 1g7 ao invés da unidade de multiplexação de fluxo de bit 1g do dispositivo de codificação de voz 11b da modificação 4 da primeira modalidade, e ainda inclui a unidade de cálculo de envelope temporal 1m e a unidade de cálculo de formato de envelope 1n do dispositivo de codificação de voz 13.

A unidade de multiplexação de fluxo de bit 1g7 multiplexa o fluxo de bit codificado calculado pela unidade de codificação de codec central 1c e as informações suplementares de SBR calculadas pela unidade de codifica-

ção de SBR 1d como a unidade de multiplexação de fluxo de bit 1g, converte o parâmetro de resistência de filtro calculado pela unidade de cálculo de parâmetro de resistência de filtro e o parâmetro de formato de envelope calculado pela unidade de cálculo de parâmetro de formato de envelope 1n nas
 5 informações suplementares de envelope temporal, multiplexa-os, e emite o fluxo de bit multiplexado (fluxo de bit multiplexado codificado) através do dispositivo de comunicação do dispositivo de codificação de voz 14.

(Modificação 4 da Quarta modalidade)

Um dispositivo codificador de voz 14a (figura 49) de uma modifi-
 10 cação 4 da quarta modalidade inclui fisicamente um CPU, um ROM, um RAM, um dispositivo de comunicação, e similares, que não são ilustrados, e o CPU controla integralmente o dispositivo codificador de voz 14a ao carregar e executar um programa de computador predeterminado armazenado em uma memória embutida do dispositivo codificador de voz 14a como o
 15 ROM no RAM. O dispositivo de comunicação do dispositivo codificador de voz 14a recebe um sinal de voz a ser codificado a partir do lado de fora do dispositivo codificador de voz 14a, e emite um fluxo de bit multiplexado codificado para o lado de fora do dispositivo codificador de voz 14a. O dispositivo codificador de voz 14a inclui a unidade de análise de predição linear 1e1 ao
 20 invés da unidade de análise de predição linear 1e do dispositivo codificador de voz 14 da quarta modalidade, e adicionalmente inclui a unidade de seleção de abertura de tempo 1p.

Um dispositivo decodificador de voz 24d (ver figura 26) da modifi-
 25 cação 4 da quarta modalidade inclui fisicamente um CPU, um ROM, um RAM, um dispositivo de comunicação, e similares, que não são ilustrados, e o CPU controla integralmente o dispositivo decodificador de voz 24d ao carregar e executar um programa de computador predeterminado (como um programa de computador para desempenhar processos ilustrados no fluxo-
 30 grama da figura 27) armazenado em uma memória embutida do dispositivo decodificador de voz 24d como o ROM no RAM. O dispositivo de comunicação do dispositivo decodificador de voz 24d recebe o fluxo de bit multiplexado codificado, e emite um sinal de voz decodificado para o lado de fora do

dispositivo decodificador de voz 24d. O dispositivo decodificador de voz 24d, conforme ilustrado na figura 26, inclui a unidade de análise de predição linear de baixa frequência 2d1, a unidade de detecção de mudança de sinal 2e1, a unidade de análise de predição linear de alta frequência 2h1, a unidade de
 5 filtro inverso de predição linear 2i1, e a unidade de filtro de predição linear 2k3 ao invés de unidade de análise de predição linear de baixa frequência 2d, a unidade de detecção de mudança de sinal 2e, a unidade de análise de predição linear de alta frequência 2h, a unidade de filtro inverso de predição linear 2i, e a unidade de filtro de predição linear 2k do dispositivo decodifica-
 10 dor de voz 24, e adicionalmente inclui a unidade de seleção de abertura de tempo 3a. A unidade de transformação de envelope temporal 2v transforma o sinal no domínio QMF obtido a partir da unidade de filtro de predição linear 2k3 através do uso de informação de envelope temporal obtida a partir da unidade de ajuste do formato de envelope 2s, conforme a unidade de trans-
 15 formação de envelope temporal 2v da terceira modalidade, da quarta modalidade, e das modificações do presente documento (processo na Etapa Sk1).
(Modificação 5 da Quarta modalidade)

Um dispositivo decodificador de voz 24e (ver figura 28) de uma modificação 5 da quarta modalidade inclui fisicamente um CPU, um ROM,
 20 um RAM, um dispositivo de comunicação, e similares, que não são ilustrados, e a CPU controla integralmente o dispositivo decodificador de voz 24e ao carregar e executar um programa de computador predeterminado (como um programa de computador para desempenhar processos ilustrados no fluxograma da figura 29) armazenado em uma memória embutida do dispositi-
 25 vo decodificador de voz 24e como o ROM no RAM. O dispositivo de comunicação do dispositivo decodificador de voz 24e recebe o fluxo de bit multiplexado codificado, e emite um sinal de voz decodificado para o lado de fora do dispositivo decodificador de voz 24e. Na modificação 5, conforme ilustrado na figura 28, o dispositivo decodificador de voz 24e omite a unidade de
 30 análise de predição linear de alta frequência 2h1 e a unidade de filtro inverso de predição linear 2i1 do dispositivo decodificador de voz 24d descritos na modificação 4 que podem ser omitidos durante toda a quarta modalidade

conforme a primeira modalidade, e inclui uma unidade de seleção de abertura de tempo 3a2 e uma unidade de transformação de envelope temporal 2v1 ao invés de unidade de seleção de abertura de tempo 3a e a unidade de transformação de envelope temporal 2v do dispositivo decodificador de voz 24d. O dispositivo decodificador de voz 24e também muda a ordem da filtra-
gem de síntese de predição linear desempenhada através da unidade de filtro de predição linear 2k3 e do processo de transformação de envelope temporal desempenhado através da unidade de transformação de envelope temporal 2v1 cuja ordem de processamento é intercambiável através da quarta modalidade.

A unidade de transformação de envelope temporal 2v1 transforma $q_{adj}(k, r)$ obtido a partir da unidade de ajuste de alta frequência 2j através do uso de $e_{adj}(r)$ obtido a partir da unidade de ajuste do formato de envelope 2s, conforme a unidade de transformação de envelope temporal 2v, e obtém um sinal $q_{envadj}(k, r)$ no domínio QMF no qual o envelope temporal é transformado. A unidade de transformação de envelope temporal 2v1 também notifica a unidade de seleção de abertura de tempo 3a2 de um parâmetro obtido quando o envelope temporal está sendo transformado, ou um parâmetro calculado ao pelo menos usar o parâmetro obtido quando o envelope temporal está sendo transformado como informação de seleção de abertura de tempo. A informação de seleção de abertura de tempo pode ser $e(r)$ da expressão (22) ou a expressão (40), ou $|e(r)|^2$ para o qual a operação de raiz quadrada não é aplicada durante o processo de cálculo. Uma pluralidade de seções de abertura de tempo (como envelopes SBR)

$$b_i \leq r < b_{i+1} \quad \text{---(49)}$$

pode também ser usado, e a expressão (24) que é o valor médio do presente documento

$$\overline{e(i)}, |\overline{e(i)}|^2 \quad \text{---(50)}$$

podem também ser usados como a informação de seleção de abertura de tempo. Note-se que:

$$\overline{|e(i)|}^2 = \frac{\sum_{r=b_i}^{b_{i+1}-1} |e(r)|^2}{b_{i+1} - b_i} \quad \text{---(51)}$$

A informação de seleção de abertura de tempo pode também ser $e_{\text{exp}}(r)$ da expressão (26) e a expressão (41), ou $|e_{\text{exp}}(r)|^2$ para o qual a operação de raiz quadrada não é aplicada durante o processo de cálculo. Uma pluralidade de seções de abertura de tempo (como envelopes SBR)

$$b_i \leq r < b_{i+1} \quad \text{---(52)}$$

5 e o valor médio do presente documento

$$\bar{e}_{\text{exp}}(i), \quad |\bar{e}_{\text{exp}}(i)|^2 \quad \text{---(53)}$$

podem também ser usados como a informação de seleção de abertura de tempo. Note-se que:

$$\bar{e}_{\text{exp}}(i) = \frac{\sum_{r=b_i}^{b_{i+1}-1} e_{\text{exp}}(r)}{b_{i+1} - b_i} \quad \text{---(54)}$$

$$|\bar{e}_{\text{exp}}(i)|^2 = \frac{\sum_{r=b_i}^{b_{i+1}-1} |e_{\text{exp}}(r)|^2}{b_{i+1} - b_i} \quad \text{---(55)}$$

10 A informação de seleção de abertura de tempo pode também ser $e_{\text{adj}}(r)$ da expressão (23), da expressão (35) ou da expressão (36), ou pode ser $|e_{\text{adj}}(r)|^2$ para o qual a operação de raiz quadrada não é aplicada durante o processo de cálculo. Uma pluralidade de seções de abertura de tempo (como envelopes SBR)

$$b_i \leq r < b_{i+1} \quad \text{---(56)}$$

e o valor médio do presente documento

$$\bar{e}_{\text{adj}}(i), \quad |\bar{e}_{\text{adj}}(i)|^2 \quad \text{---(57)}$$

15 abertura de tempo. Note-se que:

$$\bar{e}_{\text{adj}}(i) = \frac{\sum_{r=b_i}^{b_{i+1}-1} e_{\text{adj}}(r)}{b_{i+1} - b_i} \quad \text{---(58)}$$

$$|\bar{e}_{adj}(i)|^2 = \frac{\sum_{r=b_i}^{b_{i+1}-1} |e_{adj}(r)|^2}{b_{i+1} - b_i} \quad \text{---(59)}$$

A informação de seleção de abertura de tempo pode também ser $e_{adj,scaled}(r)$ da expressão (37), ou pode ser $|e_{adj,scaled}(r)|^2$ para o qual a operação de raiz quadrada não é aplicada durante o processo de cálculo. Em uma pluralidade de seções de abertura de tempo (como envelopes SBR)

$$b_i \leq r < b_{i+1} \quad \text{---(60)}$$

5 e o valor médio do presente documento

$$\bar{e}_{adj,scaled}(i), \quad |\bar{e}_{adj,scaled}(i)|^2 \quad \text{---(61)}$$

podem também ser usados como a informação de seleção de abertura de tempo. Note-se que:

$$\bar{e}_{adj,scaled}(i) = \frac{\sum_{r=b_i}^{b_{i+1}-1} e_{adj,scaled}(r)}{b_{i+1} - b_i} \quad \text{---(62)}$$

$$|\bar{e}_{adj,scaled}(i)|^2 = \frac{\sum_{r=b_i}^{b_{i+1}-1} |e_{adj,scaled}(r)|^2}{b_{i+1} - b_i} \quad \text{---(63)}$$

A informação de seleção de abertura de tempo pode também ser uma força de sinal $P_{envadj}(r)$ da abertura de tempo r do sinal de domínio QMF que corresponde aos componentes de alta frequência nos quais o envelope temporal é transformado ou um valor de amplitude de sinal do presente documento para o qual a operação de raiz quadrada é aplicada

$$\sqrt{P_{envadj}(r)} \quad \text{---(64)}$$

Em uma pluralidade de seções de abertura de tempo (como envelopes SBR)

$$b_i \leq r < b_{i+1} \quad \text{---(65)}$$

15 e o valor médio do presente documento

$$\bar{P}_{envadj}(i), \quad \sqrt{\bar{P}_{envadj}(i)} \quad \text{---(66)}$$

podem também ser usados como a informação de seleção de abertura de tempo. Note-se que:

$$P_{envadj}(r) = \sum_{k=k_x}^{k_x+M-1} |q_{envadj}(k, r)|^2 \quad \text{---(67)}$$

$$\bar{P}_{envadj}(i) = \frac{\sum_{r=b_i}^{b_{i+1}-1} P_{envadj}(r)}{b_{i+1} - b_i} \quad \text{---(68)}$$

M é um valor que representa uma faixa de frequência mais alta do que aquela da frequência de limite baixo k_x dos componentes de alta frequência gerados através da unidade de geração de alta frequência 2g, e a faixa de frequência dos componentes de alta frequência gerada através da unidade de geração de alta frequência 2g pode também ser representada como $k_x \leq k < k_x + M$.

A unidade de seleção de abertura de tempo 3a2 seleciona uma abertura de tempo na qual a filtragem de síntese de predição linear através da qual a unidade de filtro de predição linear 2k é desempenhada, ao determinar se a filtragem de síntese de predição linear é desempenhada no sinal $q_{envadj}(k, r)$ no domínio QMF dos componentes de alta frequência da abertura de tempo r na qual o envelope temporal é transformado através da unidade de transformação de envelope temporal 2v1, com base na informação de seleção de abertura de tempo transmitida a partir da unidade de transformação de envelope temporal 2v1 (processo na Etapa Sp1).

Para selecionar uma abertura de tempo na qual a filtragem de síntese de predição linear é desempenhada através da unidade de seleção de abertura de tempo 3a2 na presente modificação, pelo menos uma abertura de tempo r na qual um parâmetro $u(r)$ incluído na informação de seleção de abertura de tempo transmitida a partir da unidade de transformação de envelope temporal 2v1 é maior do que um valor predeterminado u_{Th} pode ser selecionado, ou pelo menos uma abertura de tempo r na qual $u(r)$ é igual a ou maior do que um valor predeterminado u_{Th} pode ser selecionada. $u(r)$ pode incluir pelo menos um de $e(r)$, $|e(r)|^2$, $e_{exp}(r)$, $|e_{exp}(r)|^2$, $e_{adj}(r)$, $|e_{adj}(r)|^2$, $e_{adj,scaled}(r)$, $|e_{adj,scaled}(r)|^2$, e $P_{envadj}(r)$, descrito acima, e;

$$\sqrt{P_{envadj}(r)} \quad \text{---(69)}$$

e u_{Th} pode incluir pelo menos um de;

$$\overline{e(i)}, |\overline{e(i)}|^2, e_{\text{exp}}(i),$$

$$|\overline{e_{\text{exp}}}(i)|^2, \overline{e_{\text{adj}}}(i), |\overline{e_{\text{adj}}}(i)|^2 \quad \text{---(70)}$$

$$\overline{e_{\text{adj,scaled}}}(i), |\overline{e_{\text{adj,scaled}}}(i)|^2,$$

$$\overline{P_{\text{envadj}}}(i), \sqrt{\overline{P_{\text{envadj}}}(i)}.$$

- u_{Th} pode também ser um valor médio de $u(r)$ de uma largura temporal predeterminada (como envelope SBR) incluindo a abertura de tempo r . A seleção pode também ser feita de forma que uma abertura de tempo na qual $u(r)$ alcança seu pico seja incluída. O pico de $u(r)$ pode ser calculado
- 5 conforme o cálculo do pico da força de sinal no sinal de domínio QMF dos componentes de alta frequência na modificação 4 da primeira modalidade. O estado estável e o estado transitório na modificação 4 da primeira modalidade podem ser determinados de modo similar àqueles da modificação 4 da primeira modalidade através do uso de $u(r)$, e uma abertura de tempo pode
- 10 ser selecionada com base no mesmo. O método de seleção de abertura de tempo pode ser pelo menos um dos métodos descritos acima, pode incluir pelo menos um método diferente dos descritos acima, ou pode ser a combinação dos mesmos.

(Modificação 6 da Quarta modalidade)

- 15 Um dispositivo decodificador de voz 24f (ver figura 30) de uma modificação 6 da quarta modalidade inclui fisicamente um CPU, um ROM, um RAM, um dispositivo de comunicação, e similares, que não são ilustrados, e a CPU controla integralmente o dispositivo decodificador de voz 24f ao carregar e executar um programa de computador predeterminado (como
- 20 um programa de computador para desempenhar processos ilustrados no fluxograma da figura 29) armazenado em uma memória embutida do dispositivo decodificador de voz 24f como o ROM no RAM. O dispositivo de comunicação do dispositivo decodificador de voz 24f recebe o fluxo de bit multiplexado codificado e emite um sinal de voz decodificado para fora do dispositivo
- 25 decodificador de voz 24f. Na modificação 6, conforme ilustrado na figu-

ra 30, o dispositivo decodificador de voz 24f omite a unidade de detecção de mudança de sinal 2e1, a unidade de análise de predição linear de alta frequência 2h1, e a unidade de filtro inverso de predição linear 2i1 do dispositivo decodificador de voz 24d descrito na modificação 4 que pode ser omitida
 5 através da quarta modalidade conforme a primeira modalidade, e inclui a unidade de seleção de abertura de tempo 3a2 e a unidade de transformação de envelope temporal 2v1 ao invés de unidade de seleção de abertura de tempo 3a e a unidade de transformação de envelope temporal 2v do dispositivo decodificador de voz 24d. O dispositivo decodificador de voz 24f também
 10 muda a ordem da filtragem de síntese de predição linear desempenhada através da unidade de filtro de predição linear 2k3 e o processo de transformação de envelope temporal desempenhado através da unidade de transformação de envelope temporal 2v1 cuja ordem de processamento é intercambiável durante toda a quarta modalidade.

15 A unidade de seleção de abertura de tempo 3a2 determina se a filtragem de síntese de predição linear é desempenhada através da unidade de filtro de predição linear 2k3, no sinal $q_{envadj}(k, r)$ no domínio QMF dos componentes de alta frequência da abertura de tempo r na qual o envelope temporal é transformado através da unidade de transformação de envelope temporal 2v1, com base na informação de seleção de abertura de tempo
 20 transmitida a partir da unidade de transformação de envelope temporal 2v1, seleciona a abertura de tempo na qual a filtragem de síntese de predição linear é desempenhada, e notifica, da abertura de tempo selecionada, a unidade de análise de predição linear de baixa frequência 2d1 e a unidade de
 25 filtro de predição linear 2k3.

(Modificação 7 da Quarta modalidade)

Um dispositivo codificador de voz 14b (figura 50) de uma modificação 7 da quarta modalidade inclui fisicamente um CPU, um ROM, um RAM, um dispositivo de comunicação, e similares, que não são ilustrados, e
 30 a CPU controla integralmente o dispositivo codificador de voz 14b ao carregar e executar um programa de computador predeterminado armazenado na memória embutida do dispositivo codificador de voz 14b como o ROM no

RAM. O dispositivo de comunicação do dispositivo codificador de voz 14b recebe um sinal de voz a ser codificado a partir do lado de fora o dispositivo codificador de voz 14b, e emite um fluxo de bit multiplexado codificado para o lado de fora do dispositivo codificador de voz 14b. O dispositivo codificador de voz 14b inclui uma unidade de multiplexação de fluxo de bit 1g6 e a unidade de seleção de abertura de tempo 1p1 ao invés de unidade de multiplexação de fluxo de bit 1g7 e da unidade de seleção de abertura de tempo 1p do dispositivo codificador de voz 14a da modificação 4.

A unidade de multiplexação de fluxo de bit 1g6 multiplexa o fluxo de bit codificado calculado através da unidade de codificação do codec central 1c, a Informação suplementar SBR calculada através da Unidade de codificação SBR 1d, e a informação suplementar de envelope temporal na qual o parâmetro de resistência de filtro calculado através da unidade de cálculo de parâmetro de resistência de filtro e o parâmetro de formato de envelope calculado através da unidade de cálculo de parâmetro de formato de envelope 1n são convertidos, também multiplexa a informação de seleção de abertura de tempo recebida a partir da unidade de seleção de abertura de tempo 1p1, e emite o fluxo de bit multiplexado (fluxo de bit multiplexado codificado) através do dispositivo de comunicação do dispositivo codificador de voz 14b.

Um dispositivo decodificador de voz 24g (ver figura 31) da modificação 7 da quarta modalidade inclui fisicamente um CPU, um ROM, um RAM, um dispositivo de comunicação, e similares, que não são ilustrados, e a CPU controla integralmente o dispositivo decodificador de voz 24g ao carregar e executar um programa de computador predeterminado (como um programa de computador para desempenhar processos ilustrados no fluxograma da figura 32) armazenado na memória embutida do dispositivo decodificador de voz 24g como o ROM no RAM. O dispositivo de comunicação do dispositivo decodificador de voz 24g recebe o fluxo de bit multiplexado codificado e emite um sinal de voz decodificado para fora do dispositivo decodificador de voz 24g. O dispositivo decodificador de voz 24g inclui uma unidade de separação de fluxo de bit 2a7 e a unidade de seleção de abertura de tempo 3a1 ao invés de unidade de separação de fluxo de bit 2a3 e a unidade

de seleção de abertura de tempo 3a do dispositivo decodificador de voz 24d descrito na modificação 4.

A unidade de separação de fluxo de bit 2a7 separa o fluxo de bit multiplexado abastecido através do dispositivo de comunicação do dispositivo decodificador de voz 24g para a informação suplementar de envelope temporal, a Informação suplementar SBR, e o fluxo de bit codificado, conforme a unidade de separação de fluxo de bit 2a3, e adicionalmente separa a informação de seleção de abertura de tempo.

(Modificação 8 de Quarta modalidade)

Um dispositivo decodificador de voz 24h (ver figura 33) de uma modificação 8 da quarta modalidade inclui fisicamente um CPU, um ROM, um RAM, um dispositivo de comunicação, e similares, que não são ilustrados, e a CPU controla integralmente o dispositivo decodificador de voz 24h ao carregar e executar um programa de computador predeterminado (como um programa de computador para desempenhar processos ilustrados no fluxograma da figura 34) armazenado na memória embutida do dispositivo decodificador de voz 24h como o ROM no RAM. O dispositivo de comunicação do dispositivo decodificador de voz 24h recebe o fluxo de bit multiplexado codificado e emite um sinal de voz decodificado para fora do dispositivo decodificador de voz 24h. O dispositivo decodificador de voz 24h, conforme ilustrado na figura 33, inclui a unidade de análise de predição linear de baixa frequência 2d1, a unidade de detecção de mudança de sinal 2e1, a unidade de análise de predição linear de alta frequência 2h1, a unidade de filtro inverso de predição linear 2i1, e a unidade de filtro de predição linear 2k3 ao invés de unidade de análise de predição linear de baixa frequência 2d, a unidade de detecção de mudança de sinal 2e, a unidade de análise de predição linear de alta frequência 2h, a unidade de filtro inverso de predição linear 2i, e a unidade de filtro de predição linear 2k do dispositivo decodificador de voz 24b da modificação 2, e adicionalmente inclui a unidade de seleção de abertura de tempo 3a. A unidade de ajuste de alta frequência primária 2j1 desempenha pelo menos um dos processos na etapa "Ajuste HF" em SBR em "MPEG-4 AAC", conforme a unidade de ajuste de alta frequência primária-

ria 2j1 da modificação 2 da quarta modalidade (processo na Etapa Sm1). A unidade de ajuste de alta frequência secundária 2j2 desempenha pelo menos um dos processos na etapa "Ajuste HF" em SBR em "MPEG-4 AAC", conforme a unidade de ajuste de alta frequência secundária 2j2 da modificação 2 da quarta modalidade (processo na Etapa Sm2). É preferível que o processo desempenhado através da unidade de ajuste de alta frequência secundária 2j2 seja um processo não desempenhado pela unidade de ajuste de alta frequência primária 2j1 entre os processos na etapa "Ajuste HF" em SBR em "MPEG-4 AAC".

10 (Modificação 9 da Quarta modalidade)

Um dispositivo decodificador de voz 24i (ver figura 35) da modificação 9 da quarta modalidade inclui fisicamente um CPU, um ROM, um RAM, um dispositivo de comunicação, e similares, que não são ilustrados, e a CPU controla integralmente o dispositivo decodificador de voz 24i ao carregar e executar um programa de computador predeterminado (como um programa de computador para desempenhar processos ilustrados no fluxograma da figura 36) armazenado na memória embutida do dispositivo decodificador de voz 24i como o ROM no RAM. O dispositivo de comunicação do dispositivo decodificador de voz 24i recebe o fluxo de bit multiplexado codificado e emite um sinal de voz decodificado para fora do dispositivo decodificador de voz 24i. O dispositivo decodificador de voz 24i, conforme ilustrado na figura 35, omite a unidade de análise de predição linear de alta frequência 2h1 e a unidade de filtro inverso de predição linear 2i1 do dispositivo decodificador de voz 24h da modificação 8 que pode ser omitido durante toda a quarta modalidade conforme a primeira modalidade, e inclui a unidade de transformação de envelope temporal 2v1 e a unidade de seleção de abertura de tempo 3a2 ao invés de unidade de transformação de envelope temporal 2v e a unidade de seleção de abertura de tempo 3a do dispositivo decodificador de voz 24h da modificação 8. O dispositivo decodificador de voz 24i também muda a ordem da filtragem de síntese de predição linear desempenhada através da unidade de filtro de predição linear 2k3 e o processo de transformação de envelope temporal desempenhada através da unidade de

transformação de envelope temporal 2v1 cuja ordem de processamento é intercambiável durante toda a quarta modalidade.

(Modificação 10 da Quarta modalidade)

Um dispositivo decodificador de voz 24j (ver figura 37) de uma
5 modificação 10 da quarta modalidade inclui fisicamente um CPU, um ROM,
um RAM, um dispositivo de comunicação, e similares, que não são ilustra-
dos, e a CPU controla integralmente o dispositivo decodificador de voz 24j
ao carregar e executar um programa de computador predeterminado (como
um programa de computador para desempenhar processos ilustrados no
10 fluxograma da figura 36) armazenado na memória embutida do dispositivo
decodificador de voz 24j como o ROM no RAM. O dispositivo de comunica-
ção do dispositivo decodificador de voz 24j recebe o fluxo de bit multiplexado
codificado e emite um sinal de voz decodificado para fora do dispositivo de-
codificador de voz 24j. O dispositivo decodificador de voz 24j, conforme ilus-
15 trado na figura 37, omite a unidade de detecção de mudança de sinal 2e1, a
unidade de análise de predição linear de alta frequência 2h1, e a unidade de
filtro inverso de predição linear 2i1 do dispositivo decodificador de voz 24h
da modificação 8 que pode ser omitida durante toda a quarta modalidade
conforme a primeira modalidade, e inclui a unidade de transformação de en-
20 velope temporal 2v1 e a unidade de seleção de abertura de tempo 3a2 ao
invés de unidade de transformação de envelope temporal 2v e a unidade de
seleção de abertura de tempo 3a do dispositivo decodificador de voz 24h da
modificação 8. O dispositivo decodificador de voz 24j também muda a ordem
da filtragem de síntese de predição linear desempenhada pela unidade de
25 filtro de predição linear 2k3 e o processo de transformação de envelope
temporal desempenhada através da unidade de transformação de envelope
temporal 2v1 cuja ordem de processamento é intercambiável durante toda a
quarta modalidade.

(Modificação 11 da Quarta modalidade)

30 Um dispositivo decodificador de voz 24k (ver figura 38) de uma
modificação 11 da quarta modalidade inclui fisicamente um CPU, um ROM,
um RAM, um dispositivo de comunicação, e similares, que não são ilustra-

dos, e a CPU controla integralmente o dispositivo decodificador de voz 24k ao carregar e executar um programa de computador predeterminado (como um programa de computador para desempenhar processos ilustrados no fluxograma da figura 39) armazenado na memória embutida do dispositivo decodificador de voz 24k como o ROM no RAM. O dispositivo de comunicação do dispositivo decodificador de voz 24k recebe o fluxo de bit multiplexado codificado e emite um sinal de voz decodificado para fora do dispositivo decodificador de voz 24k. O dispositivo decodificador de voz 24k, conforme ilustrado na figura 38, inclui a unidade de separação de fluxo de bit 2a7 e a unidade de seleção de abertura de tempo 3a1 ao invés de unidade de separação de fluxo de bit 2a3 e a unidade de seleção de abertura de tempo 3a do dispositivo decodificador de voz 24h da modificação 8.

(Modificação 12 da Quarta modalidade)

Um dispositivo decodificador de voz 24q (ver figura 40) de uma modificação 12 da quarta modalidade inclui fisicamente um CPU, um ROM, um RAM, um dispositivo de comunicação, e similares, que não são ilustrados, e a CPU controla integralmente o dispositivo decodificador de voz 24q ao carregar e executar um programa de computador predeterminado (como um programa de computador para desempenhar processos ilustrados no fluxograma da figura 41) armazenado na memória embutida do dispositivo decodificador de voz 24q como o ROM no RAM. O dispositivo de comunicação do dispositivo decodificador de voz 24q recebe o fluxo de bit multiplexado codificado e emite um sinal de voz decodificado para fora do dispositivo decodificador de voz 24q. O dispositivo decodificador de voz 24q, conforme ilustrado na figura 40, inclui a unidade de análise de predição linear de baixa frequência 2d1, a unidade de detecção de mudança de sinal 2e1, a unidade de análise de predição linear de alta frequência 2h1, a unidade de filtro inverso de predição linear 2i1, e unidades de ajuste de componente de sinal individual 2z4, 2z5, e 2z6 (unidades de ajuste de componente de sinal individual correspondem aos meios de transformação de envelope temporal) ao invés de unidade de análise de predição linear de baixa frequência 2d, a unidade de detecção de mudança de sinal 2e, a unidade de análise de predi-

ção linear de alta frequência 2h, a unidade de filtro inverso de predição linear 2i, e as unidades de ajuste de componente de sinal individual 2z1, 2z2, e 2z3 do dispositivo decodificador de voz 24c da modificação 3, e adicionalmente inclui a unidade de seleção de abertura de tempo 3a.

- 5 Pelo menos uma das unidades de ajuste de componente de sinal individual 2z4, 2z5, e 2z6 desempenham processamento no sinal de domínio QMF da abertura de tempo selecionada, para o componente de sinal incluso na saída da unidade de ajuste de alta frequência primária, conforme as unidades de ajuste de componente de sinal individual 2z1, 2z2, e 2z3,
- 10 com base no resultado da seleção transmitida a partir da unidade de seleção de abertura de tempo 3a (processo na Etapa Sn1). É preferível que o processo que usa a informação de seleção de abertura de tempo inclua pelo menos um processo que inclui a filtragem de síntese de predição linear na direção de frequência, entre os processos das unidades de ajuste de com-
- 15 ponente de sinal individual 2z1, 2z2, e 2z3 descritas na modificação 3 da quarta modalidade.

- Os processos desempenhados através das unidades de ajuste de componente de sinal individual 2z4, 2z5, e 2z6 podem ser os mesmos que os processos desempenhados através das unidades de ajuste de com-
- 20 ponente de sinal individual 2z1, 2z2, e 2z3 descritas na modificação 3 da quarta modalidade, mas as unidades de ajuste de componente de sinal individual 2z4, 2z5, e 2z6 podem transformar o envelope temporal de cada um de uma pluralidade de componentes de sinal inclusos na saída da unidade de ajuste de alta frequência primária através de diferentes métodos (se to-
- 25 das as unidades de ajuste de componente de sinal individual 2z4, 2z5, e 2z6 não desempenham processamento com base no resultado da seleção transmitida a partir da unidade de seleção de abertura de tempo 3a, será o mesmo que a modificação 3 da quarta modalidade da presente invenção).

- Todos os resultados da seleção da abertura de tempo transmiti-
- 30 da para as unidades de ajuste de componente de sinal individual 2z4, 2z5, e 2z6 a partir da unidade de seleção de abertura de tempo 3a não precisam

ser os mesmos, e todo ou uma parte do presente documento pode ser diferente.

Na figura 40, o resultado da seleção da abertura de tempo é transmitida para as unidades de ajuste de componente de sinal individual 2z4, 2z5, e 2z6 a partir de uma unidade de seleção de abertura de tempo 3a. Entretanto, é possível incluir uma pluralidade de unidades de seleção de abertura de tempo para notificar, dos diferentes resultados da seleção da abertura de tempo, cada ou uma parte das unidades de ajuste de componente de sinal individual 2z4, 2z5, e 2z6. Neste momento, a unidade de seleção de abertura de tempo em relação à unidade de ajuste de componente de sinal individual entre as unidades de ajuste de componente de sinal individual 2z4, 2z5, e 2z6 que desempenham o processo 4 (o processo de multiplicar o coeficiente de ganho através de cada resolução de sub-banda QMF é desempenhada no sinal de entrada através do uso do envelope temporal obtido a partir da unidade de ajuste do formato de envelope 2s conforme a unidade de transformação de envelope temporal 2v, e então a filtragem de síntese de predição linear na direção de frequência é também desempenhada no sinal de saída através do uso do coeficiente de predição linear recebido a partir da unidade de ajuste de resistência de filtro 2f conforme a unidade de filtro de predição linear 2k) descrito na modificação 3 da quarta modalidade pode selecionar a abertura de tempo através do uso da informação de seleção de abertura de tempo abastecido a partir da unidade de transformação de envelope temporal.

(Modificação 13 da Quarta modalidade)

Um dispositivo decodificador de voz 24m (ver figura 42) de uma modificação 13 da quarta modalidade inclui fisicamente um CPU, um ROM, um RAM, um dispositivo de comunicação, e similares, que não são ilustrados, e a CPU controla integralmente o dispositivo decodificador de voz 24m ao carregar e executar um programa de computador predeterminado (como um programa de computador para desempenhar processos ilustrados no fluxograma da figura 43) armazenado na memória embutida do dispositivo decodificador de voz 24m como o ROM no RAM. O dispositivo de comunica-

ção do dispositivo decodificador de voz 24m recebe o fluxo de bit multiplexado codificado e emite um sinal de voz decodificado para fora do dispositivo decodificador de voz 24m. O dispositivo decodificador de voz 24m, conforme ilustrado na figura 42, inclui a unidade de separação de fluxo de bit 2a7 e a

5 unidade de seleção de abertura de tempo 3a1 ao invés de unidade de separação de fluxo de bit 2a3 e a unidade de seleção de abertura de tempo 3a do dispositivo decodificador de voz 24q da modificação 12.

(Modificação 14 da Quarta modalidade)

Um dispositivo decodificador de voz 24n (não ilustrado) de uma

10 modificação 14 da quarta modalidade inclui fisicamente um CPU, um ROM, um RAM, um dispositivo de comunicação, e similares, que não são ilustrados, e a CPU controla integralmente o dispositivo decodificador de voz 24n ao carregar e executar um programa de computador predeterminado armazenado na memória embutida do dispositivo decodificador de voz 24n como

15 o ROM no RAM. O dispositivo de comunicação do dispositivo decodificador de voz 24n recebe o fluxo de bit multiplexado codificado e emite um sinal de voz decodificado para fora do dispositivo decodificador de voz 24n. O dispositivo decodificador de voz 24n inclui funcionalmente a unidade de análise de

predição linear de baixa frequência 2d1, a unidade de detecção de mudança de sinal 2e1, a unidade de análise de predição linear de alta frequência 2h1,

20 a unidade de filtro inverso de predição linear 2i1, e a unidade de filtro de predição linear 2k3 ao invés de unidade de análise de predição linear de baixa frequência 2d, a unidade de detecção de mudança de sinal 2e, a unidade de análise de predição linear de alta frequência 2h, a unidade de filtro inverso de predição linear 2i,

25 e a unidade de filtro de predição linear 2k do dispositivo decodificador de voz 24a da modificação 1, e adicionalmente inclui a unidade de seleção de abertura de tempo 3a.

(Modificação 15 da Quarta modalidade)

Um dispositivo decodificador de voz 24p (não ilustrado) de uma

30 modificação 15 da quarta modalidade inclui fisicamente um CPU, um ROM, um RAM, um dispositivo de comunicação, e similares, que não são ilustrados, e a CPU controla integralmente o dispositivo decodificador de voz 24p

ao carregar e executar um programa de computador predeterminado armazenado na memória embutida do dispositivo decodificador de voz 24p como o ROM no RAM. O dispositivo de comunicação do dispositivo decodificador de voz 24p recebe o fluxo de bit multiplexado codificado e emite um sinal de voz decodificado para fora do dispositivo decodificador de voz 24p. O dispositivo decodificador de voz 24p funcionalmente inclui a unidade de seleção de abertura de tempo 3a1 ao invés de unidade de seleção de abertura de tempo 3a do dispositivo decodificador de voz 24n da modificação 14. O dispositivo decodificador de voz 24p também inclui uma unidade de separação de fluxo de bit 2a8 (não ilustrado) ao invés de unidade de separação de fluxo de bit 2a4.

A unidade de separação de fluxo de bit 2a8 separa o fluxo de bit multiplexado na informação suplementar SBR e o fluxo de bit codificado conforme a unidade de separação de fluxo de bit 2a4, e adicionalmente na formação de seleção de abertura de tempo.

Aplicabilidade Industrial

A presente invenção fornece uma técnica aplicável à técnica de extensão de banda no domínio de frequência representado por SBR, e para reduzir a ocorrência de pré-eco e pós-eco e melhora a qualidade subjetiva do sinal decodificado sem aumentar significativamente a taxa de bit.

Lista de Signos de Referência

	11, 11a, 11b, 11c, 12, 12a,	
	12b, 13, 14, 14a, 14b	dispositivo de codificação de voz
	1a	unidade de conversão de frequência
25	1b	unidade de conversão inversa de frequência
	1c	unidade de codificação de codec central
	1d	unidade de codificação de SBR
	1e, 1e1	unidade de análise de previsão linear
	1f	unidade de cálculo de parâmetro de resistência de
30		filtro
	1f1	unidade de cálculo de parâmetro de resistência de
		filtro

	1g, 1g1, 1g2, 1g3, 1g4, 1g5, 1g6, 1g7	unidade de multiplexação de fluxo de bit
	1h	unidade de conversão inversa de alta frequência
	1i	unidade de cálculo de força de período curto
5	1j	unidade de decimação de coeficiente de previsão linear
	1k	unidade de quantização de coeficiente de previsão
	1m	unidade de cálculo de envelope temporal
	1n	unidade de cálculo de parâmetro de formato de en-
10	velope 1p, 1p1	unidade de seleção de abertura de tempo
	21, 22, 23, 24, 24b, 24c	dispositivo de decodificação de voz
	2a, 2a1, 2a2, 2a3, 15 2a5, 2a6, 2a7	unidade de separação de fluxo de bit
	2b	unidade de decodificação de codec central
	2c	unidade de conversão de frequência
	2d, 2d1	unidade de análise de previsão linear de baixa frequência
20	2e, 2e1	unidade de detecção de sinal
	2f	unidade de ajuste de resistência de filtro
	2g	unidade de geração de alta frequência
	2h, 2h1	unidade de análise de previsão linear de alta frequência
25	2i, 2i1	unidade de filtro inverso de previsão linear
	2j, 2j1, 2j2, 2j3, 2j4	unidade de ajuste de alta frequência
	2k, 2k1, 2k2, 2k3	unidade de filtro de previsão linear
	2m	unidade de adição de coeficiente
	2n	unidade de conversão inversa de frequência
30	2p, 2p1	unidade de interpolação/extrapolação de coeficiente de previsão linear

	2r	unidade de cálculo de envelope temporal de baixa frequência
	2s	unidade de ajuste de formato de envelope
	2t	unidade de cálculo de envelope temporal de alta frequência
5	quência	
	2u	unidade de suavização de envelope temporal
	2v, 2v1	unidade de transformação de envelope temporal
	2w	unidade de conversão de informações suplementares
10	2z1, 2z2, 2z3, 2z4, 2z5, 2z6	unidade de ajuste de componente de sinal individual
	3a, 3a1, 3a2	unidade de seleção de abertura de tempo

REIVINDICAÇÕES

1. Dispositivo de decodificação de voz para decodificar um sinal de voz codificado, o dispositivo de decodificação de voz compreende:

5 meios de separação de fluxo de bit para separar um fluxo de bit que inclui o sinal de voz codificado em um fluxo de bit codificado e informações suplementares de envelope temporal, o fluxo de bit recebido de fora do dispositivo de decodificação de voz;

10 meios de decodificação central para decodificar o fluxo de bit codificado separado pelos meios de separação de fluxo de bit para obter um componente de baixa frequência;

meios de transformada de frequência para transformar o componente de baixa frequência obtido pelos meios de decodificação central para o domínio de frequência;

15 meios de geração de alta frequência para gerar um componente de alta frequência copiando o componente de baixa frequência transformado para o domínio de frequência pelos meios de transformada de frequência a partir de uma banda de baixa frequência para uma banda de alta frequência;

20 meios de ajuste de alta frequência para ajustar o componente de alta frequência gerado pelos meios de geração de alta frequência para gerar um componente de alta frequência ajustado;

25 meios de análise de envelope temporal de baixa frequência para analisar o componente de baixa frequência transformado para o domínio de frequência pelos meios de conversão de frequência para obter as informações de envelope temporal, os meios de análise de envelope temporal de baixa frequência obtendo a informação de envelope temporal pela obtenção de potência de cada amostra de sub-banda QMF do componente de baixa frequência transformado para o domínio da frequência pelos meios de transformada de frequência;

30 meios de conversão de informação suplementar para converter a informação suplementar de envelope temporal em um parâmetro para ajustar a informação de envelope temporal;

meios de ajuste de envelope temporal para ajustar as informações de envelope temporal obtidas pelos meios de análise de envelope temporal de baixa frequência para gerar informação de envelope temporal ajustada, os meios de ajuste de envelope temporal usando o parâmetro no ajuste da informação de envelope temporal; e

meios de formatação de envelope temporal para formatar um envelope temporal do componente de alta frequência ajustado, usando a informação de envelope temporal ajustada.

2. Dispositivo de decodificação de voz para decodificar um sinal de voz codificado, o dispositivo de decodificação de voz compreende:

meios de decodificação central para decodificar um fluxo de bit que inclui o sinal de voz codificado para obter um componente de baixa frequência, o fluxo de bit recebido de fora do dispositivo de decodificação de voz;

meios de transformada de frequência para transformar o componente de baixa frequência obtido pelo meio de decodificação central para o domínio de frequência;

meios de geração de alta frequência para gerar um componente de alta frequência copiando o componente de baixa frequência transformado para o domínio de frequência pelos meios de transformada de frequência de uma banda de baixa frequência para uma banda de alta frequência;

meios de ajuste de alta frequência para ajustar o componente de alta frequência gerado pelos meios de geração de alta frequência para gerar um componente de alta frequência ajustado;

meios de análise de envelope temporal de baixa frequência para analisar o componente de baixa frequência transformado para o domínio de frequência pelos meios de transformada de frequência para obter as informações de envelope temporal, os meios de análise de envelope temporal de baixa frequência obtendo a informação de envelope temporal pela obtenção de potência de cada amostra de sub-banda QMF do componente de baixa frequência transformado para o domínio da frequência pelos meios de transformada de frequência;

meios de geração de informações suplementares de envelope temporal para analisar o fluxo de bit para gerar um parâmetro para ajustar a informação de envelope temporal;

5 meios de ajuste de envelope temporal para ajustar as informações de envelope temporal obtidas pelos meios de análise de envelope temporal de baixa frequência para gerar informação de envelope temporal ajustado, os meios de ajuste de envelope temporal usando o parâmetro no ajuste da informação de envelope temporal;

10 meios de formatação de envelope temporal para formatar um envelope temporal do componente de alta frequência ajustado usando a informação de envelope temporal ajustada.

3. Dispositivo de decodificação de voz, de acordo com a reivindicação 1 ou 2, em que os meios de análise de envelope temporal de baixa frequência obtêm a informação de envelope temporal pela normalização da
15 potência de cada amostra de sub-banda de QMF pelo uso de um potência média em um segmento de tempo de envelope SBR.

4. Método de decodificação de voz que usa um dispositivo de decodificação de voz para decodificar um sinal de voz codificado, o método de decodificação de voz compreende as etapas de:

20 uma etapa de separação de fluxo de bit em que o dispositivo de decodificação de voz separa um fluxo de bit que inclui o sinal de voz codificado em um fluxo de bit codificado e informações suplementares de envelope temporal, o fluxo de bit recebido de fora do dispositivo de decodificação de voz;

25 uma etapa de decodificação central em que o dispositivo de decodificação de voz obtém um componente de baixa frequência decodificando o fluxo de bit codificado separado na etapa de separação de fluxo de bit;

30 uma etapa de transformada de frequência em que o dispositivo de decodificação de voz transforma o componente de baixa frequência obtido na etapa de decodificação central para o domínio de frequência;

uma etapa de geração de alta frequência em que o dispositivo de decodificação de voz gera um componente de alta frequência copiando o

componente de baixa frequência transformado para o domínio de frequência na etapa de transformada de frequência de uma banda de baixa frequência para uma banda de alta frequência;

- 5 uma etapa de ajuste de alta frequência na qual o dispositivo de decodificação de voz ajusta o componente de alta frequência na etapa de geração de alta frequência para gerar um componente de frequência alta ajustada;

- 10 uma etapa de análise de envelope temporal de baixa frequência em que o dispositivo de decodificação de voz obtém as informações de envelope temporal analisando o componente de baixa frequência transformado para o domínio de frequência na etapa de transformada de frequência, em que as informações de envelope temporal são obtidas pela obtenção da potência de cada amostra de sub-banda QMF do componente de baixa frequência transformado para o domínio da frequência pela etapa de transformada de frequência;
- 15

uma etapa de conversão de informação suplementar na qual o dispositivo de decodificação de voz converte a informação de envelope temporal suplementar em um parâmetro para ajustar a informação de envelope temporal;

- 20 uma etapa de ajuste de envelope temporal em que o dispositivo de decodificação de voz ajusta as informações de envelope temporal obtidas na etapa de análise de envelope temporal de baixa frequência para gerar uma informação de envelope temporal ajustada em que o parâmetro é utilizado no ajuste da informação de envelope temporal; e

- 25 uma etapa de formatação de envelope temporal em que o dispositivo de decodificação de voz formata um envelope temporal do componente de alta frequência ajustado usando as informações de envelope temporal ajustadas.

- 30 5. Método de decodificação de voz usando um dispositivo de decodificação de voz para decodificar um sinal de voz codificado, o método de decodificação de voz compreendendo as etapas de:

uma etapa de decodificação central na qual o dispositivo de decodificação de voz decodifica um fluxo de bits que inclui um sinal de voz codificado para obter um componente de baixa frequência, o fluxo de bit recebido a partir do lado de fora do dispositivo de decodificação de voz;

- 5 uma etapa de transformada de frequência na qual o dispositivo de decodificação de voz transforma o componente de frequência baixa obtido na etapa de decodificação central para o domínio da frequência;

- uma etapa de geração de alta frequência na qual o dispositivo de codificação de voz gera um componente de alta frequência pela cópia do
10 componente de baixa frequência transformado para o domínio da frequência na etapa de transformada de frequência a partir de uma banda de baixa frequência em uma banda de alta frequência;

- uma etapa de ajuste de alta frequência na qual o dispositivo de decodificação de voz ajusta o componente de frequência alta gerado na etapa de geração de alta frequência para gerar um componente de alta frequência ajustado;
15

- uma etapa de análise de envelope temporal de baixa frequência na qual o dispositivo de decodificação de voz obtém uma informação de envelope temporal pela análise do componente de frequência baixa transformado para o domínio da frequência na etapa de transformada de frequência,
20 em que as informações de envelope temporal são obtidas pela obtenção da potência de cada amostra de sub-banda QMF do componente de baixa frequência transformado para o domínio da frequência pela etapa de transformada de frequência;

- 25 uma etapa de geração de informação suplementar de envelope temporal na qual o dispositivo de decodificação de voz analisa o fluxo de bit para gerar um parâmetro para ajustar a informação de envelope temporal;

- uma etapa de ajuste de envelope temporal na qual o dispositivo de decodificação de voz ajusta a informação de envelope temporal obtida na
30 etapa de análise de envelope temporal de baixa frequência para gerar uma informação de envelope temporal ajustada, em que o parâmetro é utilizado no ajuste da informação de envelope temporal; e

uma etapa de formatação de envelope temporal na qual o dispositivo de decodificação de voz formata um envelope temporal do componente de alta frequência ajustado, usando a informação de envelope temporal ajustada.

- 5 6. Programa de decodificação de voz para decodificar um sinal de voz codificado que faz com que um dispositivo de computador funcione como:

 meio de separação de fluxo de bit para separar um fluxo de bit que inclui o sinal de voz codificado em um fluxo de bit codificado e informações suplementares de envelope temporal, o fluxo de bit recebido de fora do
10 programa de decodificação de voz;

 meio de decodificação central para decodificar o fluxo de bit codificado separado pelos meios de separação de fluxo de bit para obter um componente de baixa frequência;

- 15 meio de transformada de frequência para transformar o componente de baixa frequência obtido pelo meio de decodificação central para um domínio de frequência;

 meio de geração de alta frequência para gerar um componente de alta frequência copiando o componente de baixa frequência transformado
20 no domínio de frequência pelo meio de transformada de frequência de uma banda de baixa frequência para uma banda de alta frequência;

 meio de ajuste de alta frequência para ajustar o componente de alta frequência gerado pelo meio de geração de alta frequência para gerar um componente de alta frequência ajustado;

- 25 meio de análise de envelope temporal de baixa frequência para analisar o componente de baixa frequência transformado para o domínio de frequência pelo meio de transformada de frequência para obter informações de envelope temporal, o meio de análise de envelope temporal de baixa frequência obtendo as informações de envelope temporal pela obtenção da
30 potência de cada amostra de sub-banda QMF do componente de baixa frequência transformado para o domínio da frequência pelo meio de transformada de frequência;

meio de conversão de informação suplementar para converter a informação suplementar de envelope temporal em um parâmetro para ajustar a informação de envelope temporal;

5 meio de ajuste de envelope temporal para ajustar as informações de envelope temporal obtidas pelo meio de análise de envelope temporal de baixa frequência para gerar informação de envelope temporal ajustada, o meio de ajuste de envelope temporal usando o parâmetro no ajuste da informação de envelope temporal; e

10 meio de formatação de envelope temporal para formatar um envelope temporal do componente de alta frequência ajustado usando informação de envelope temporal ajustada.

7. Programa de decodificação de voz para decodificar um sinal de voz codificado fazendo com que um dispositivo de computação funcione
15 como:

meio de decodificação central para decodificar um fluxo de bit que inclui o sinal de voz decodificado para obter um componente de baixa frequência, o fluxo de bit recebido a partir do lado de fora do dispositivo de decodificação de voz;

20 meio de transformada de frequência para transformar o componente de baixa frequência obtido pelo meio de decodificação central para o domínio de frequência;

meio de geração de alta frequência para gerar um componente de alta frequência copiando o componente de baixa frequência transformado
25 para o domínio de frequência pelo meio de transformada de frequência de uma banda de baixa frequência para uma banda de alta frequência;

meio de ajuste de alta frequência para ajustar o componente de alta frequência gerado pelo meio de geração de alta frequência para gerar um componente de alta frequência ajustado;

30 meio de análise de envelope temporal de baixa frequência para analisar o componente de baixa frequência transformado para o domínio de frequência pelo meio de transformada de frequência para obter as informa-

ções de envelope temporal, o meio de análise de envelope temporal de baixa frequência obtendo as informações de envelope temporal pela obtenção da potência de cada amostra de sub-banda QMF do componente de baixa frequência transformado para o domínio da frequência pelo meio de trans-

5 formada de frequência;

meio de geração de informações suplementares de envelope temporal para analisar o fluxo de bit para gerar um parâmetro para ajustar a informação de envelope temporal;

meio de ajuste de envelope temporal para ajustar as informações de envelope temporal obtidas pelo meio de análise de envelope temporal de baixa frequência para gerar informação de envelope temporal ajustado, o meio de ajuste de envelope temporal usando o parâmetro no ajuste da informação de envelope temporal;

10

meio de formatação de envelope temporal para formatar um envelope temporal do componente de alta frequência ajustado usando a informação de envelope temporal ajustada.

15

Fig. 1

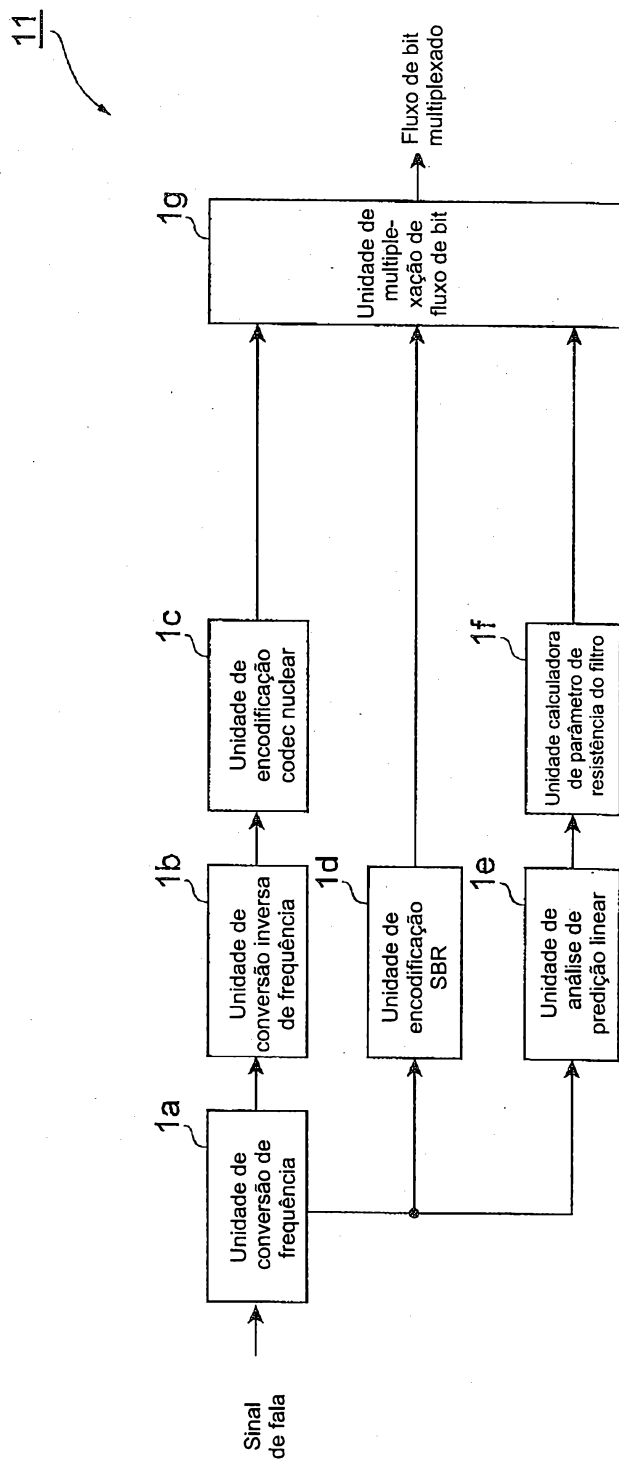


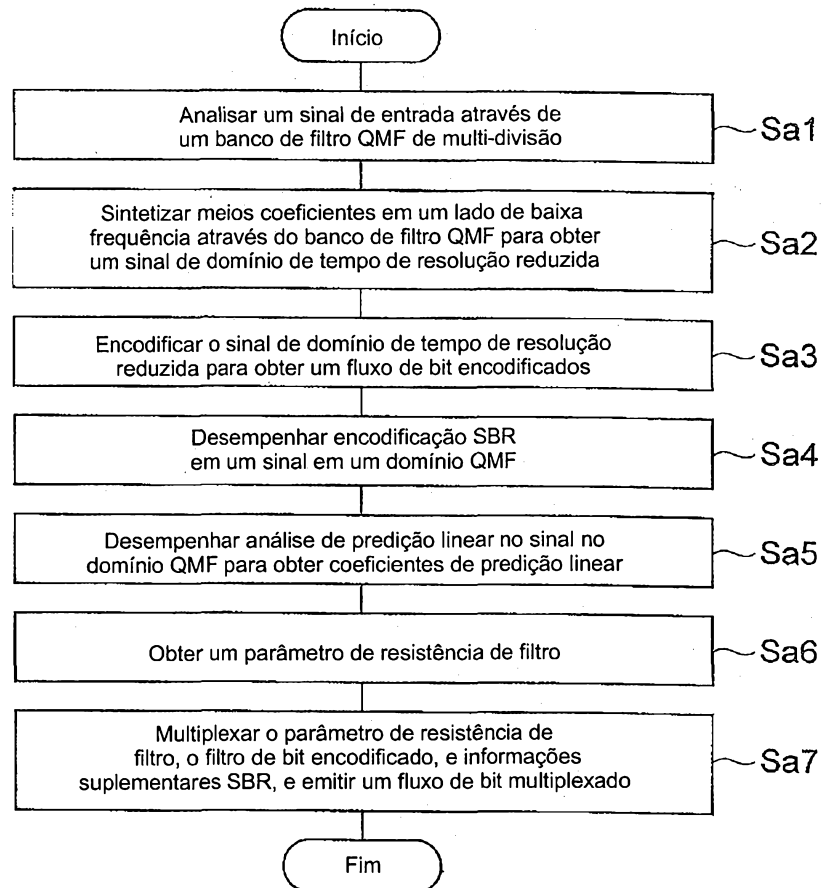
Fig.2

Fig.3

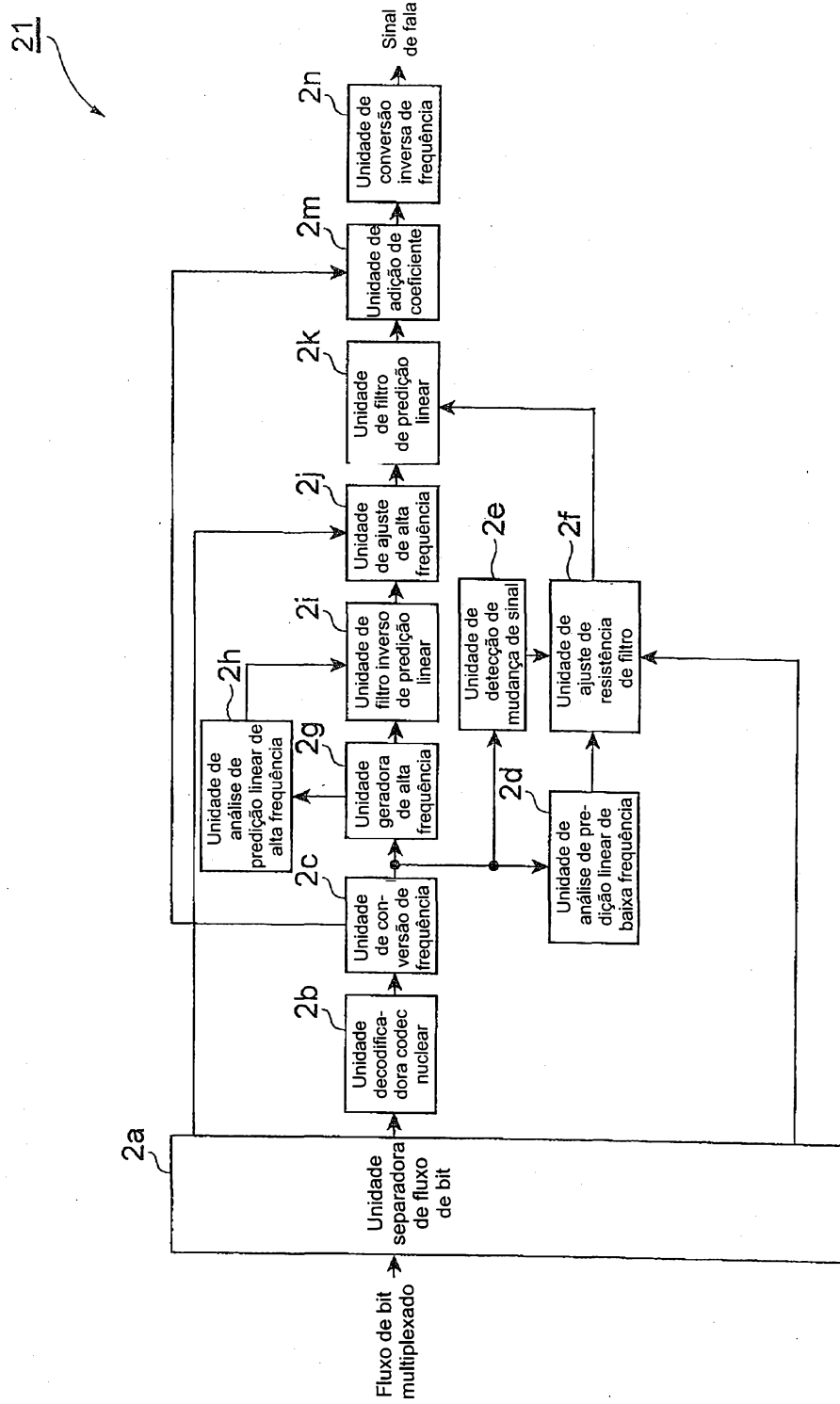


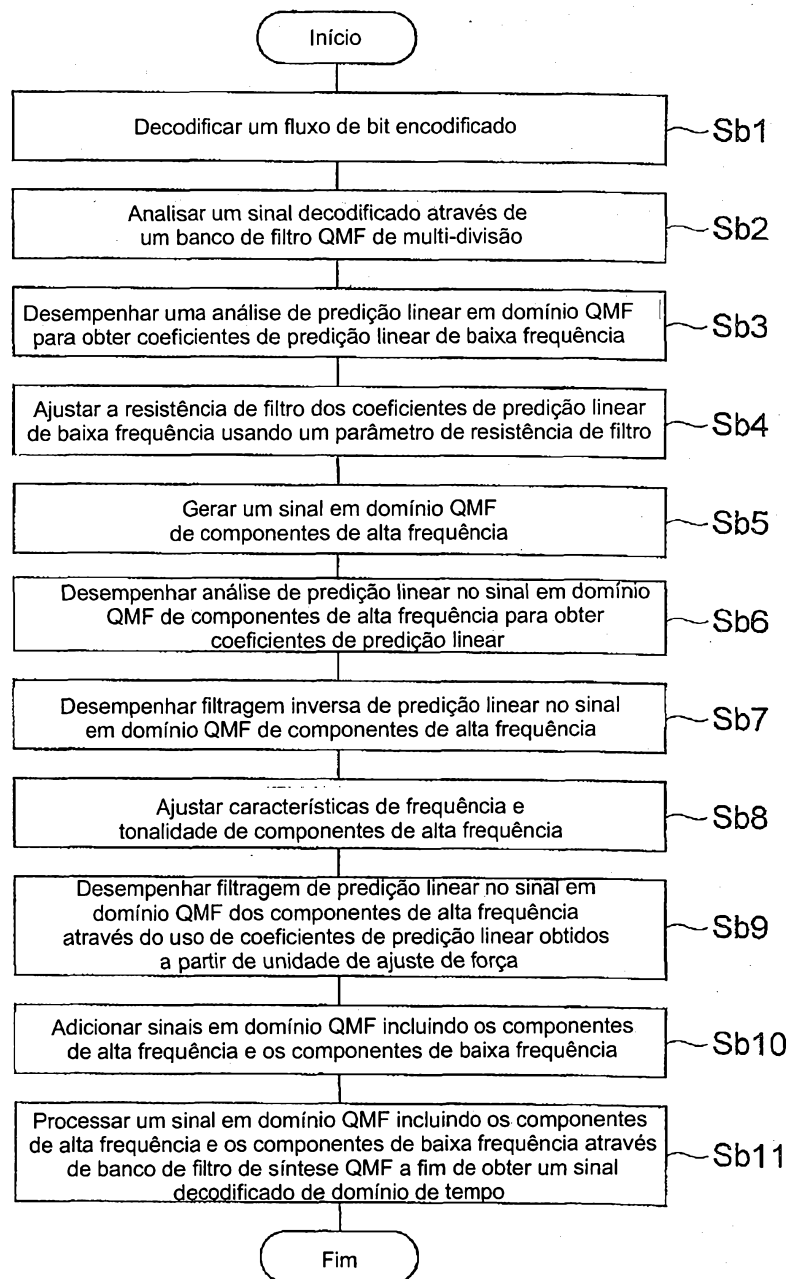
Fig.4

Fig. 5

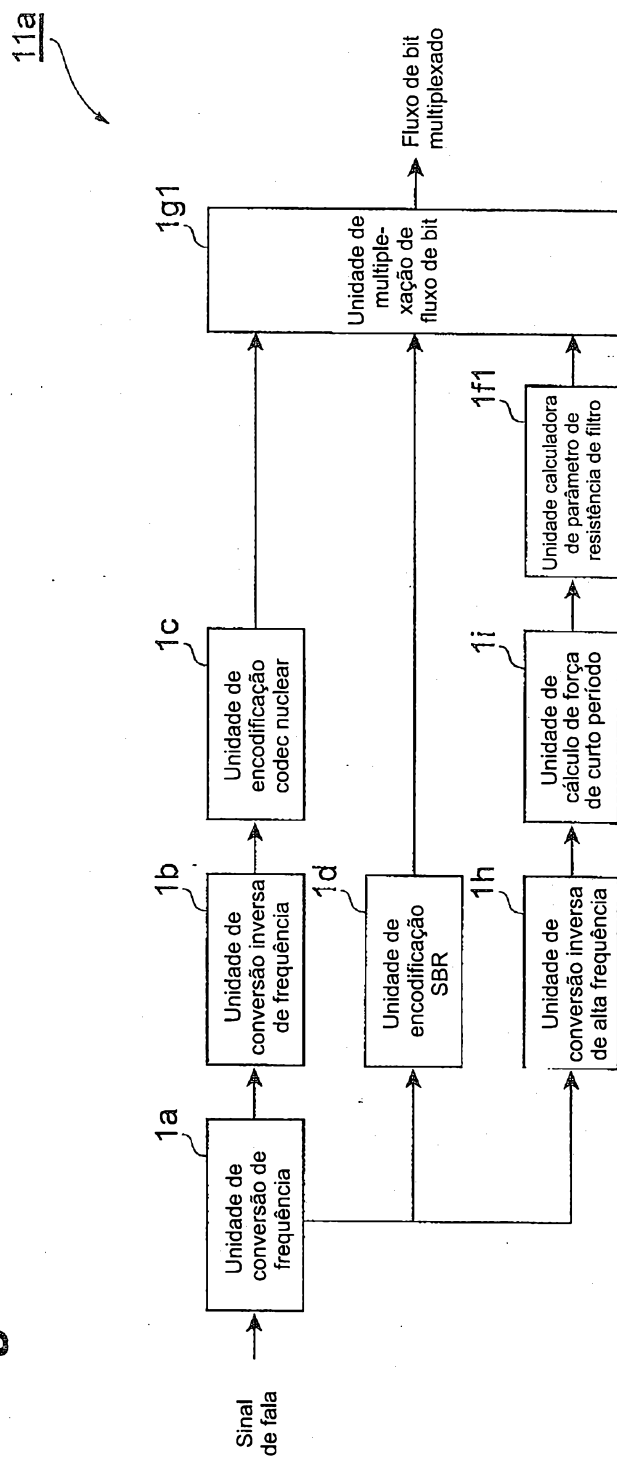


Fig. 6

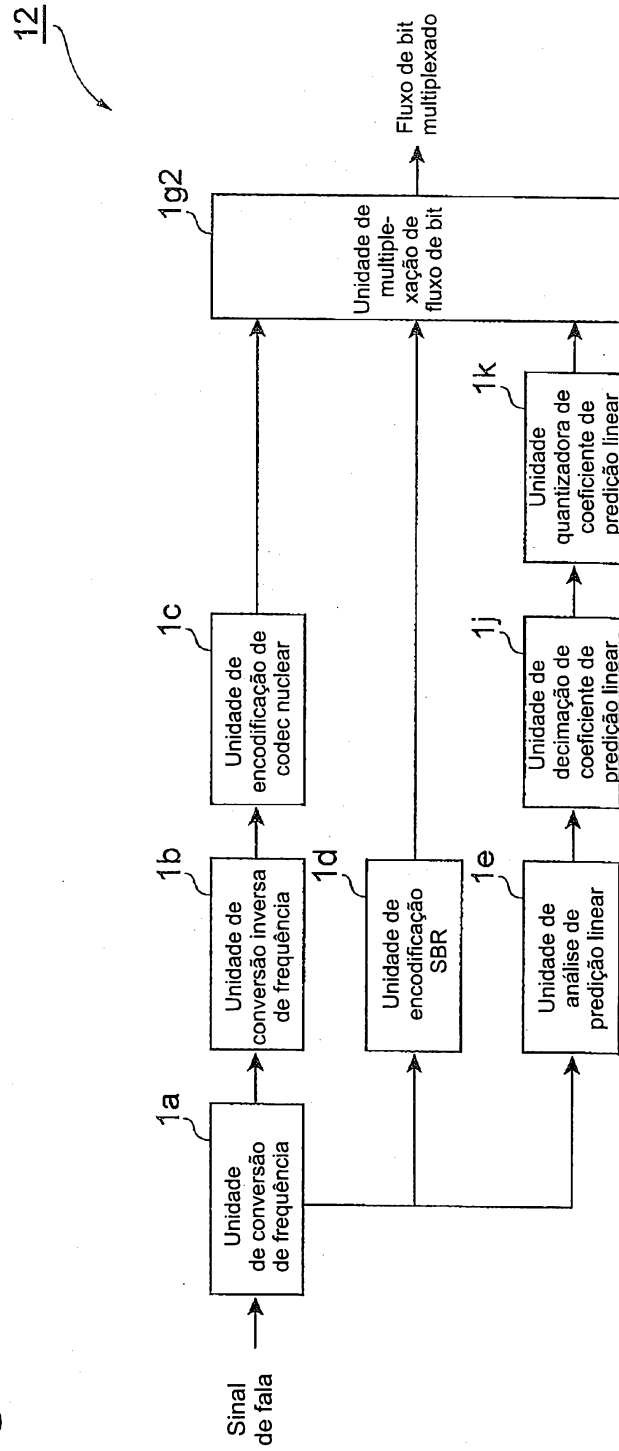


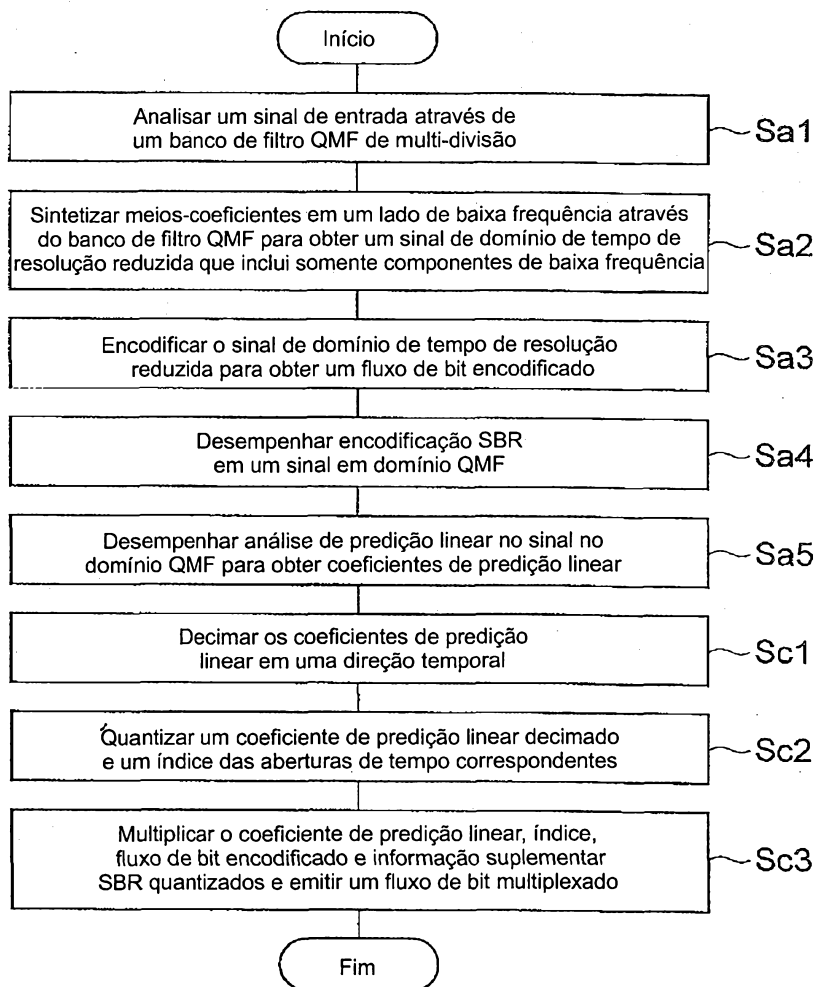
Fig.7

Fig.8

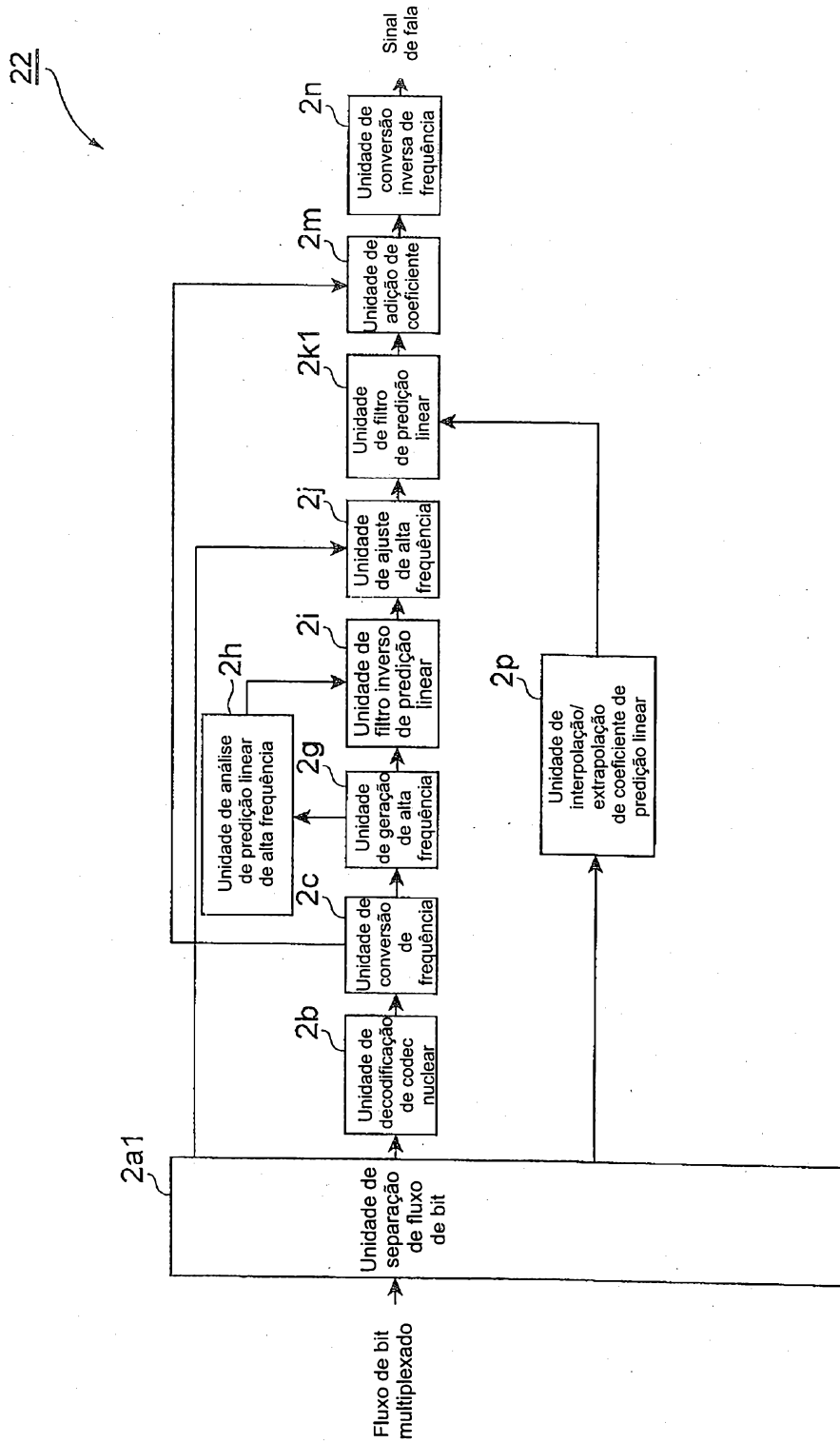


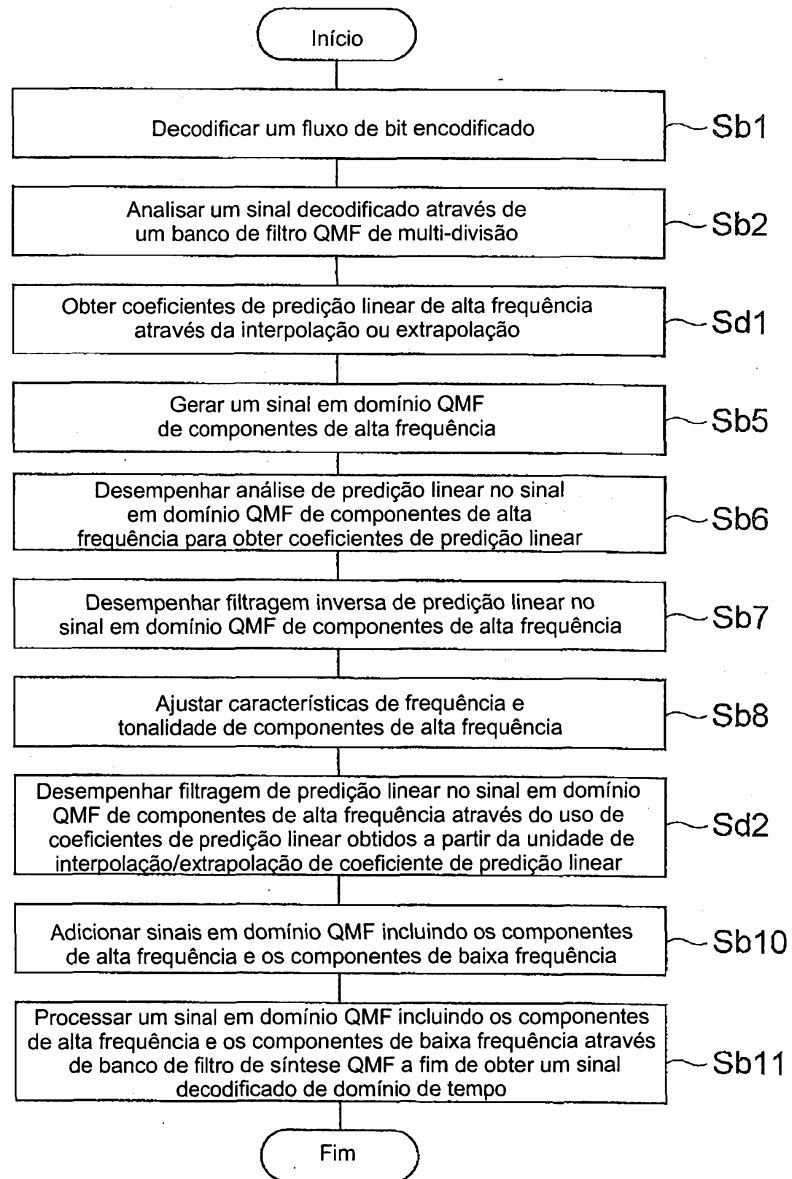
Fig.9

Fig. 10

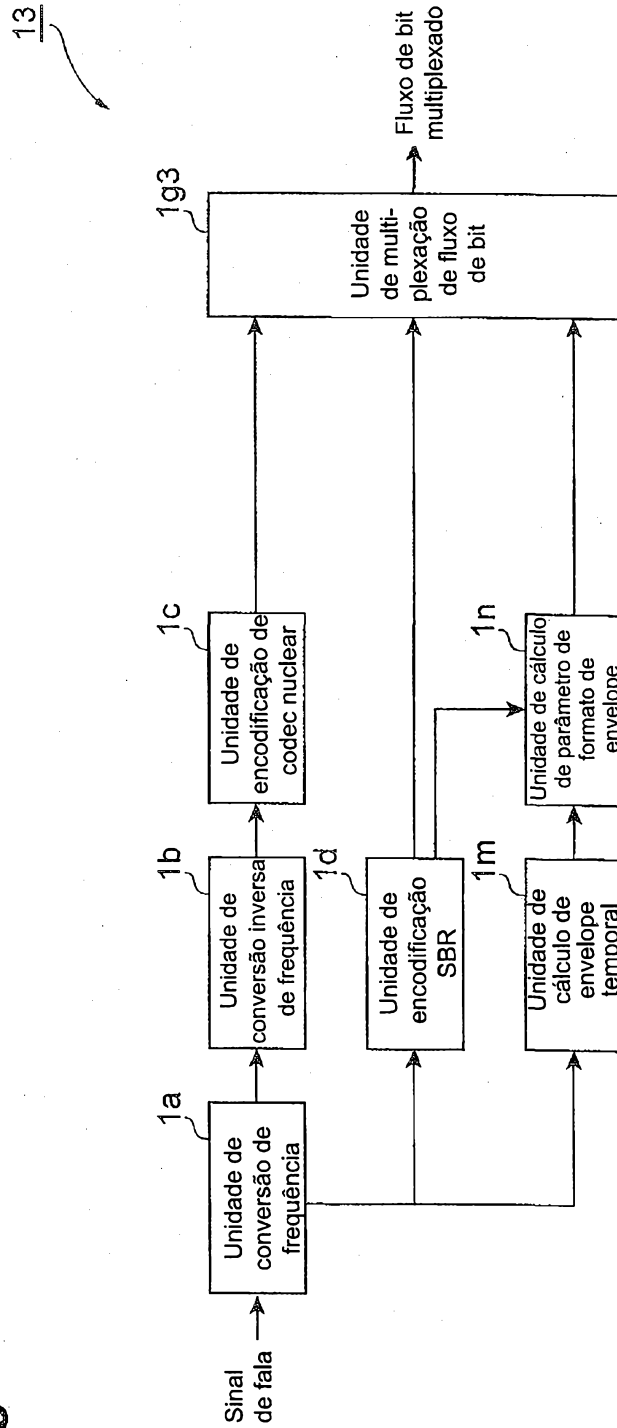


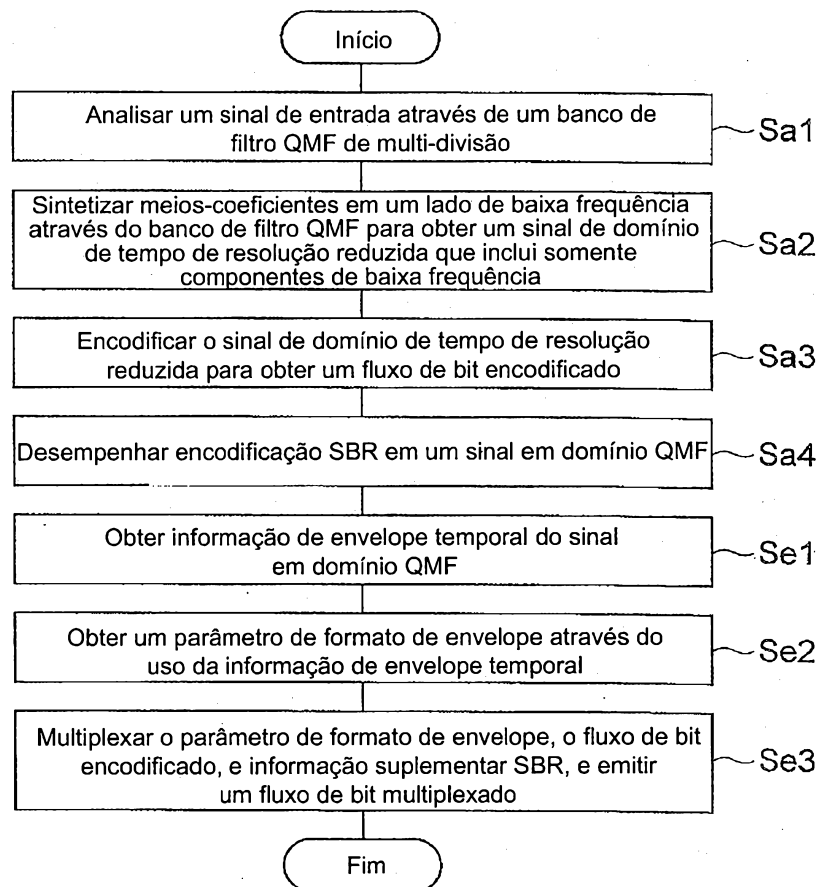
Fig.11

Fig.12

23

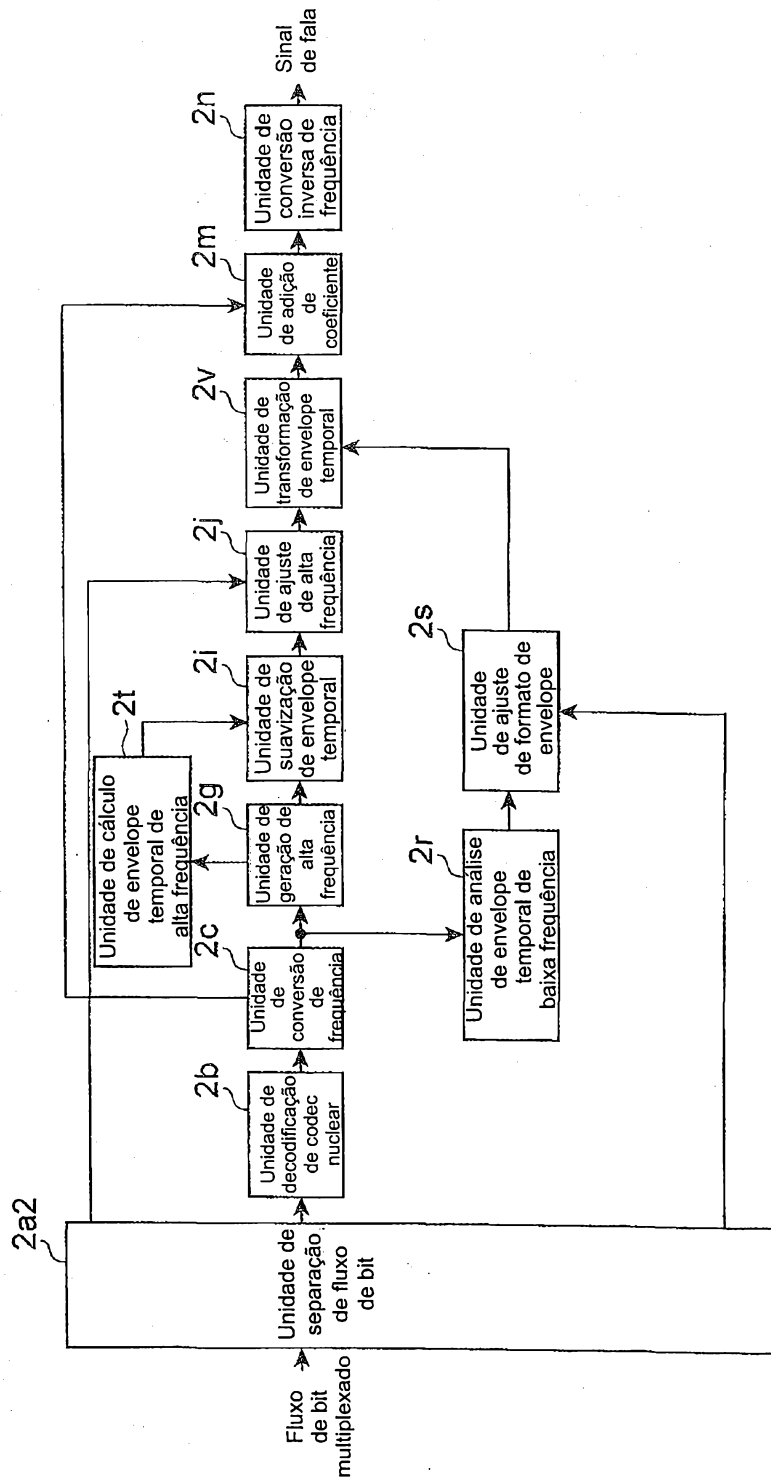


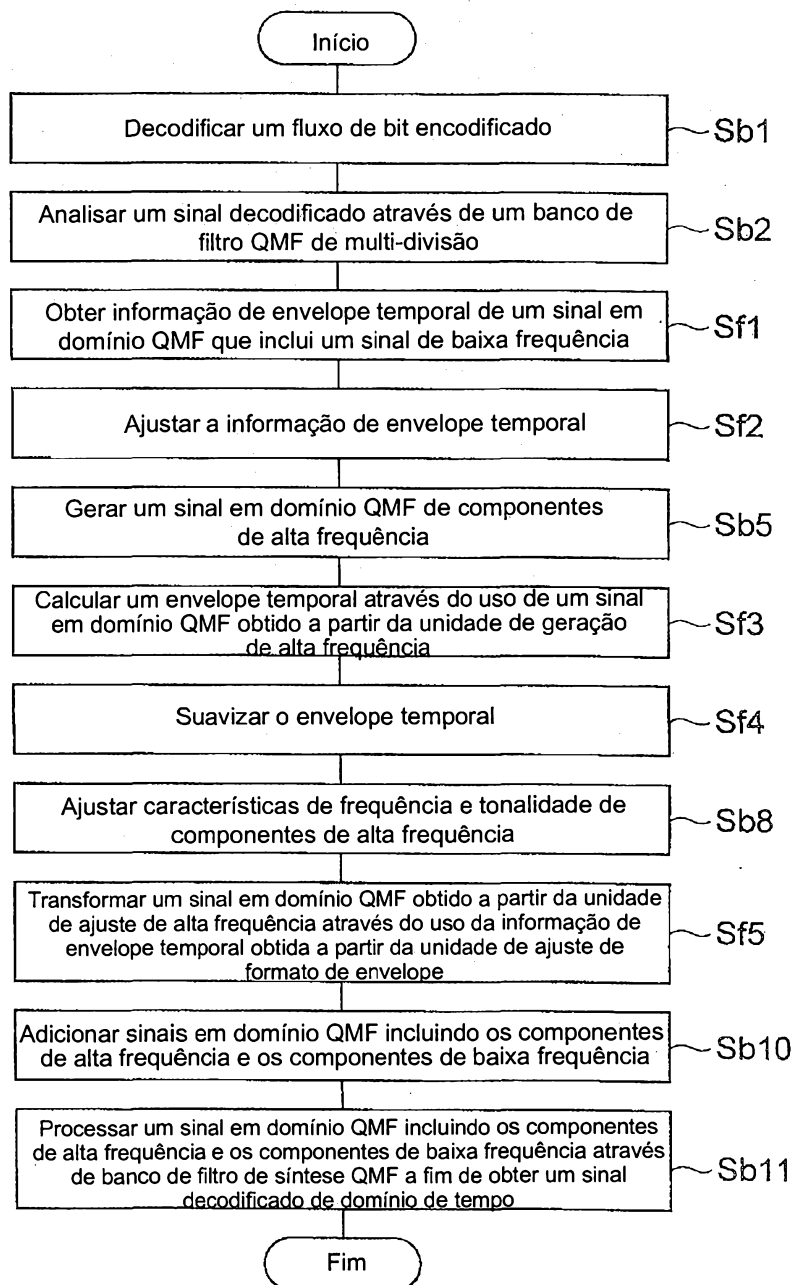
Fig.13

Fig. 14

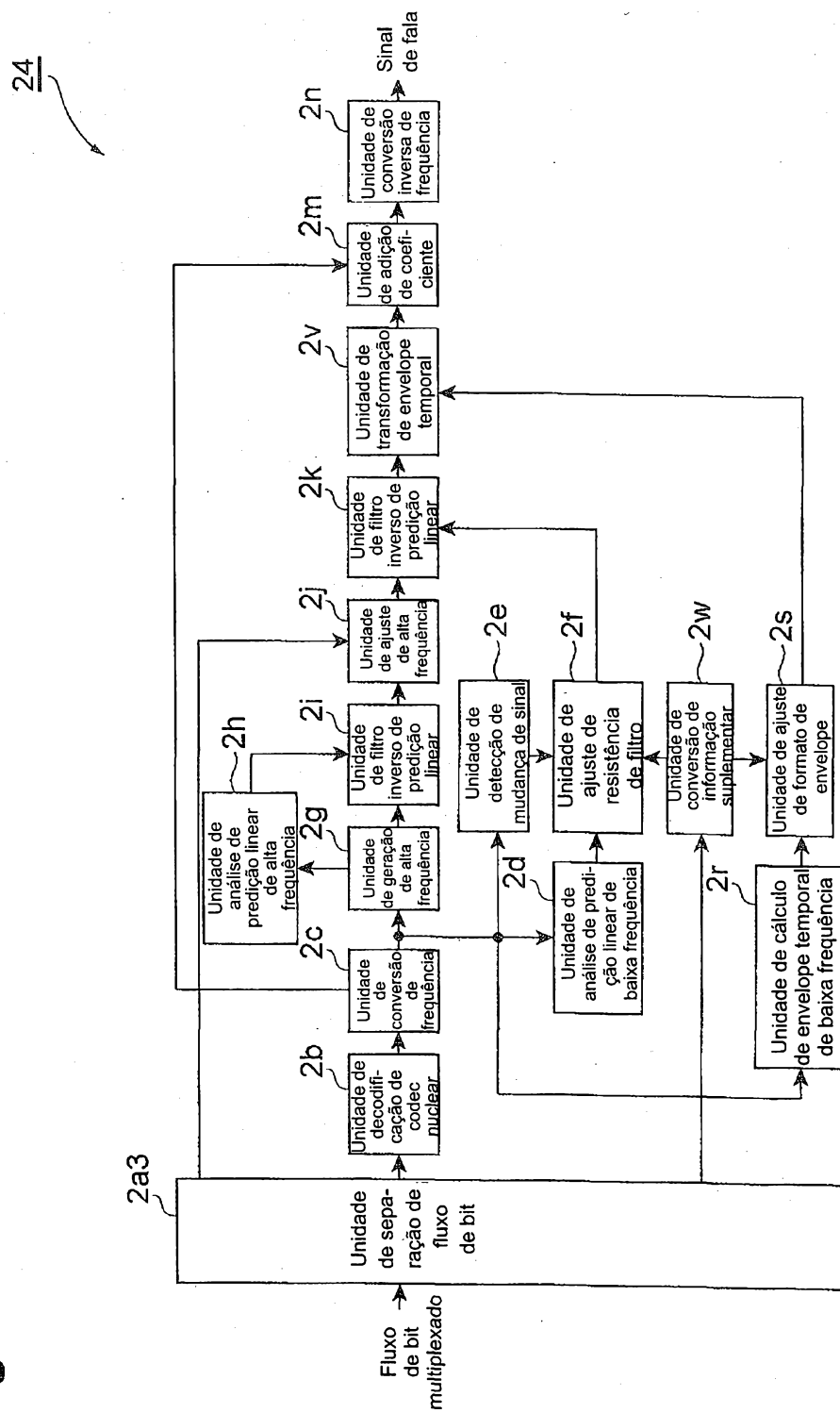


Fig.15

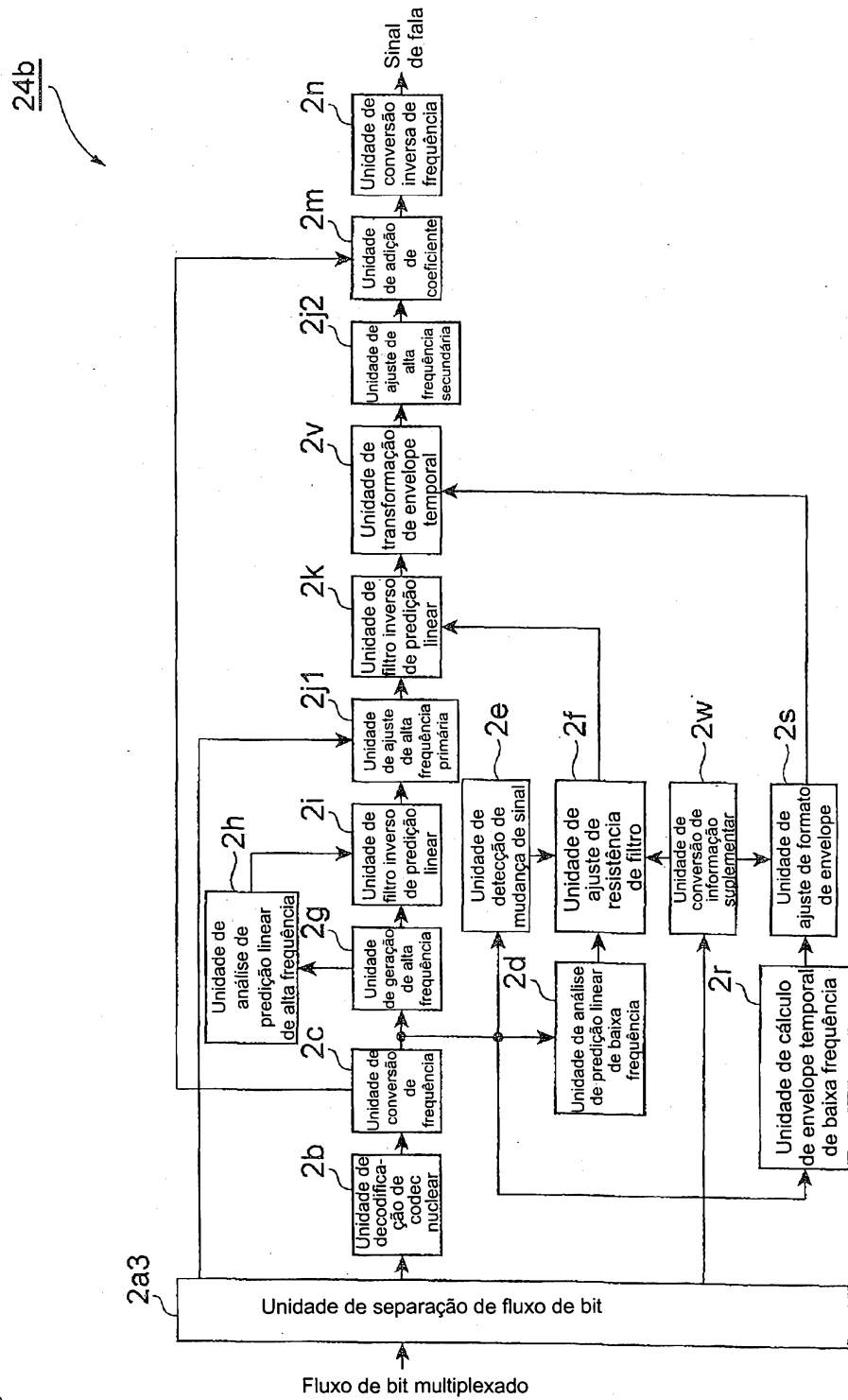


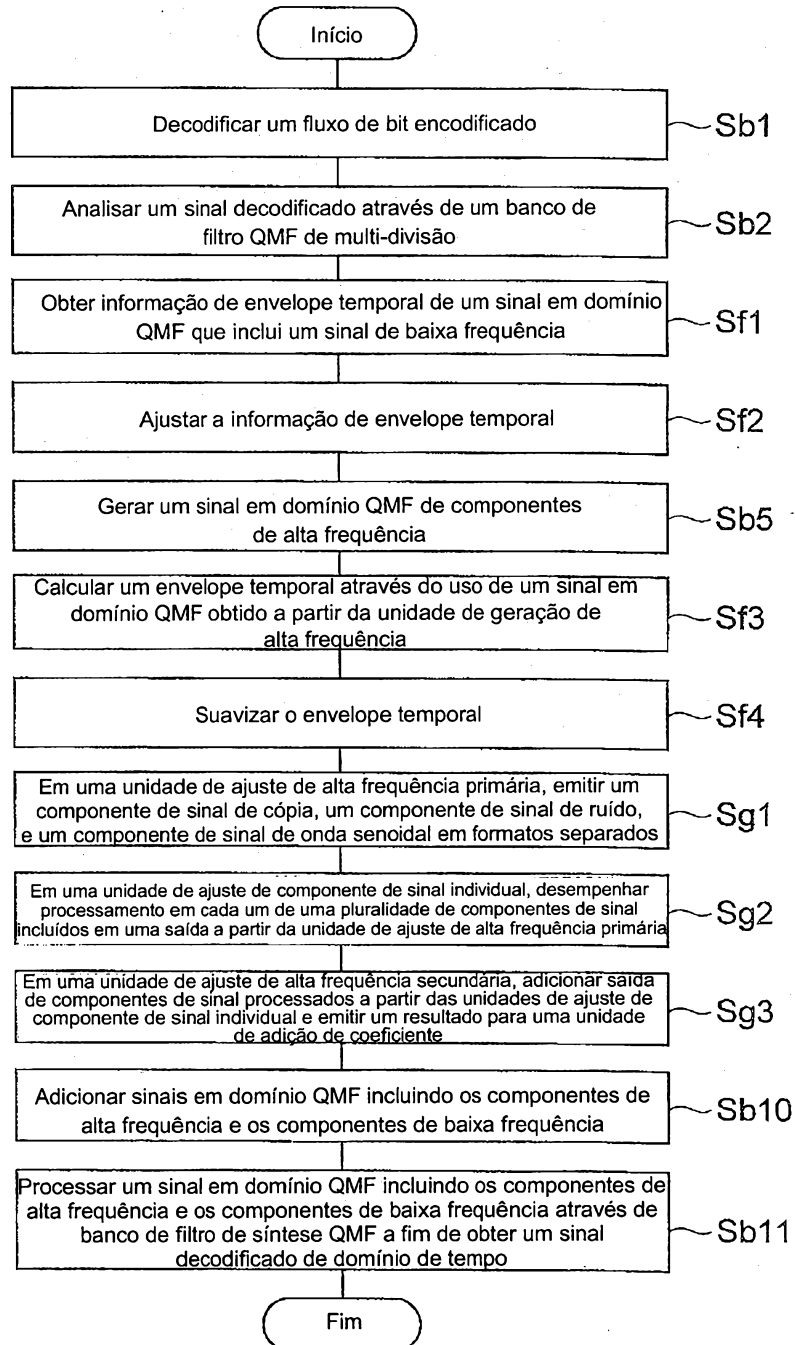
Fig.17

Fig. 18

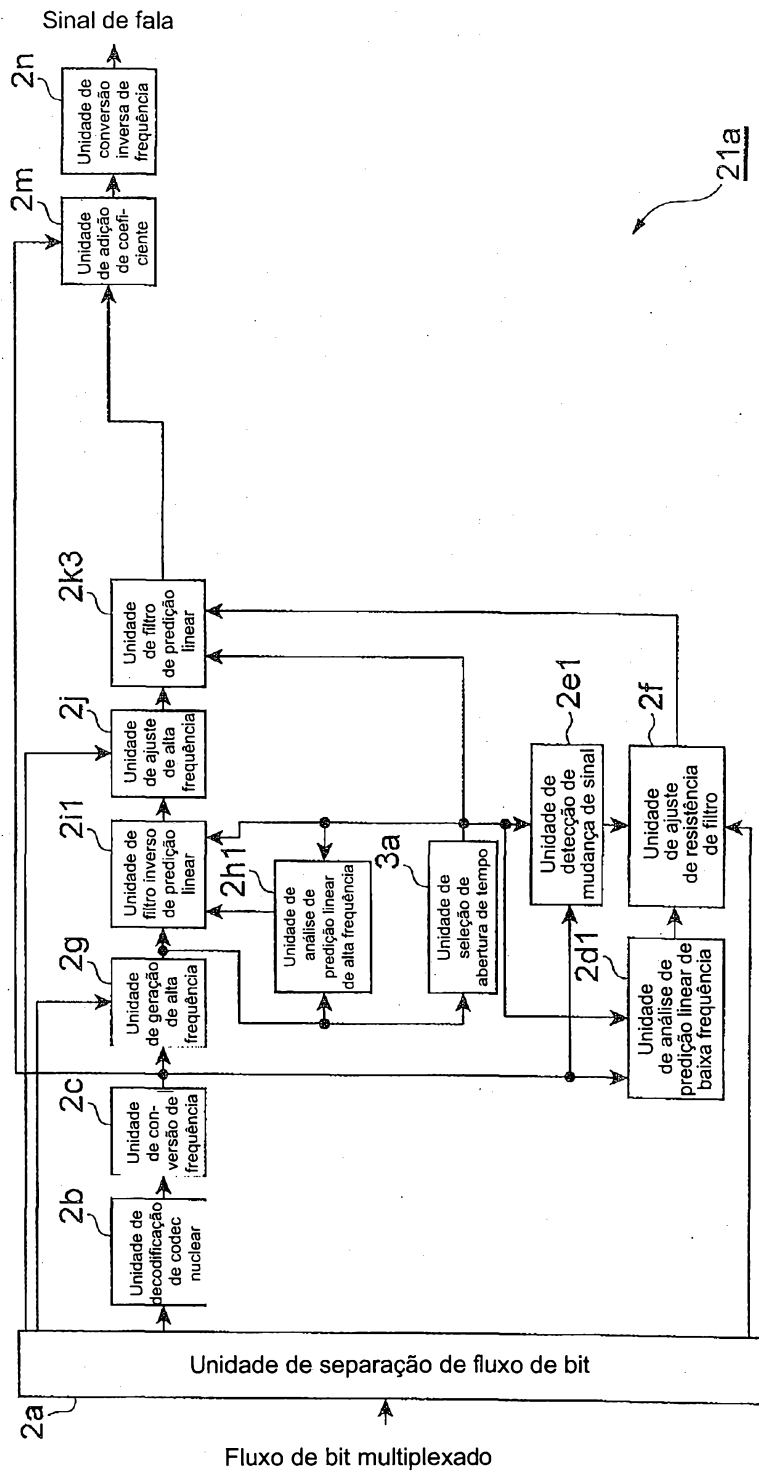


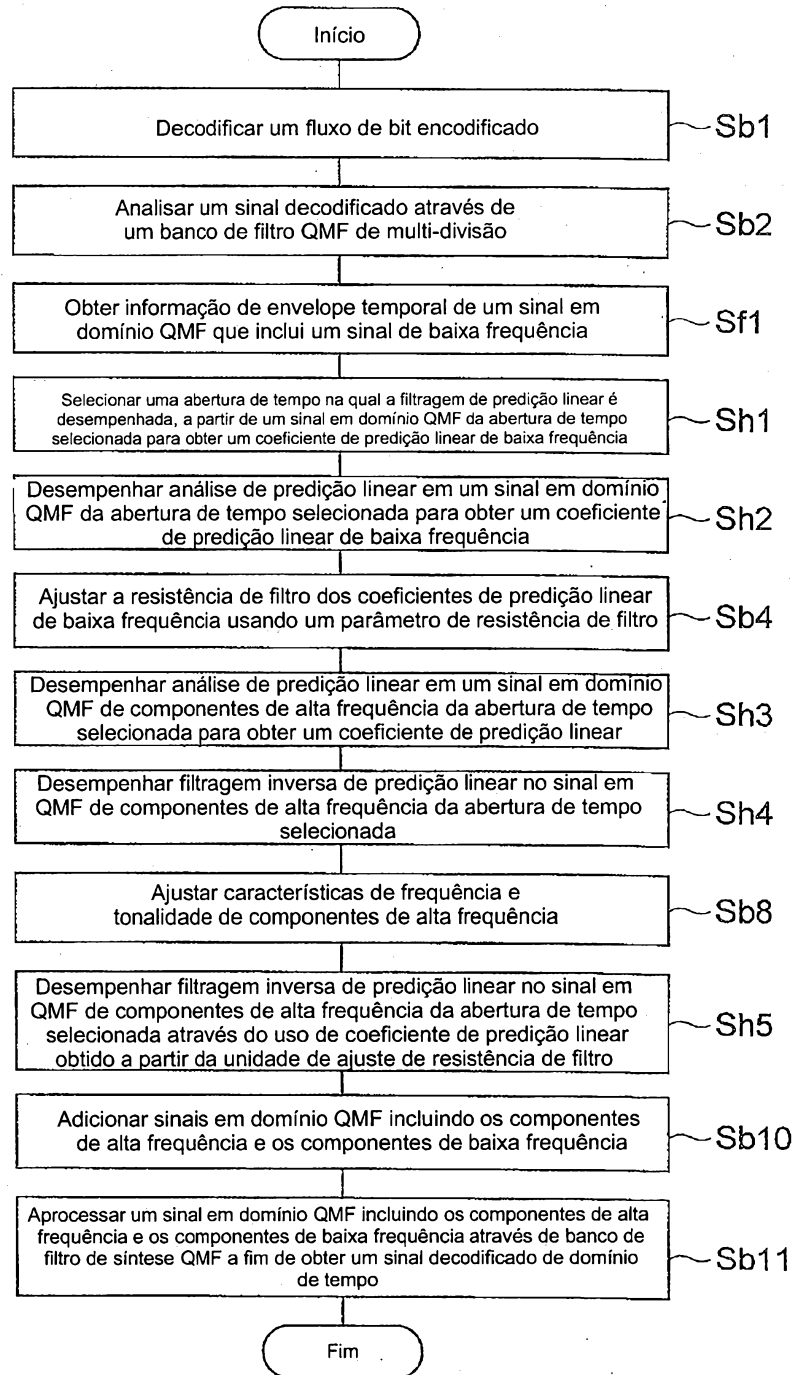
Fig.19

Fig. 20

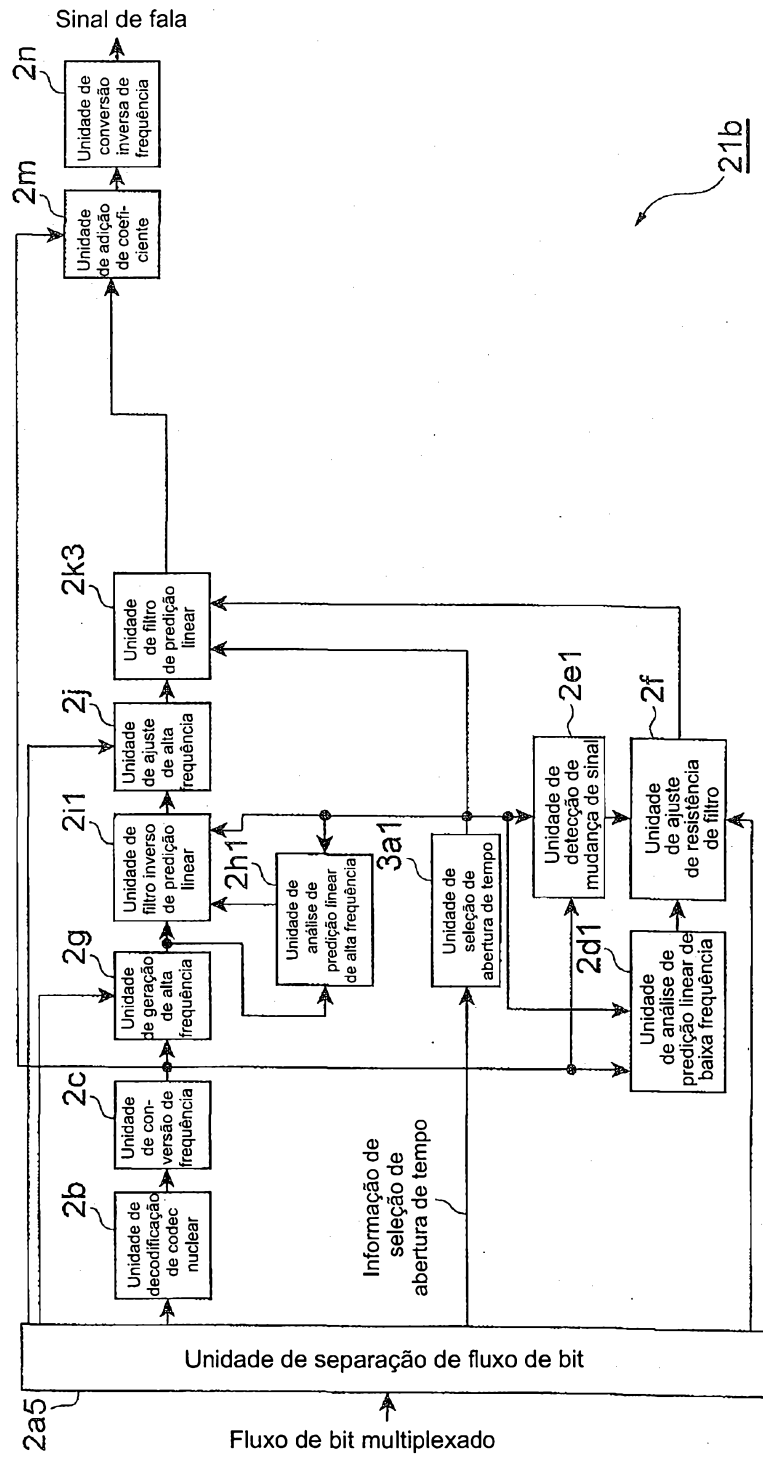


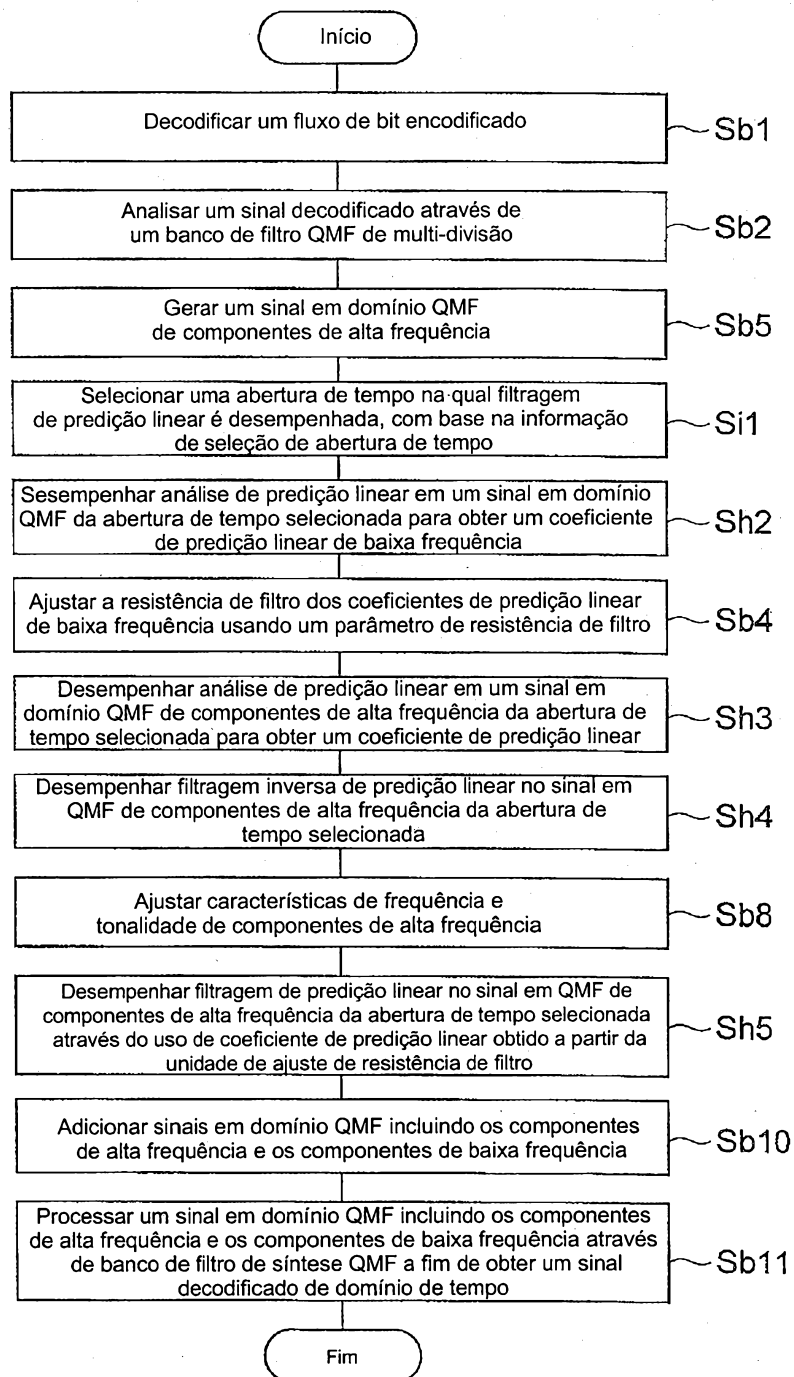
Fig.21

Fig. 22

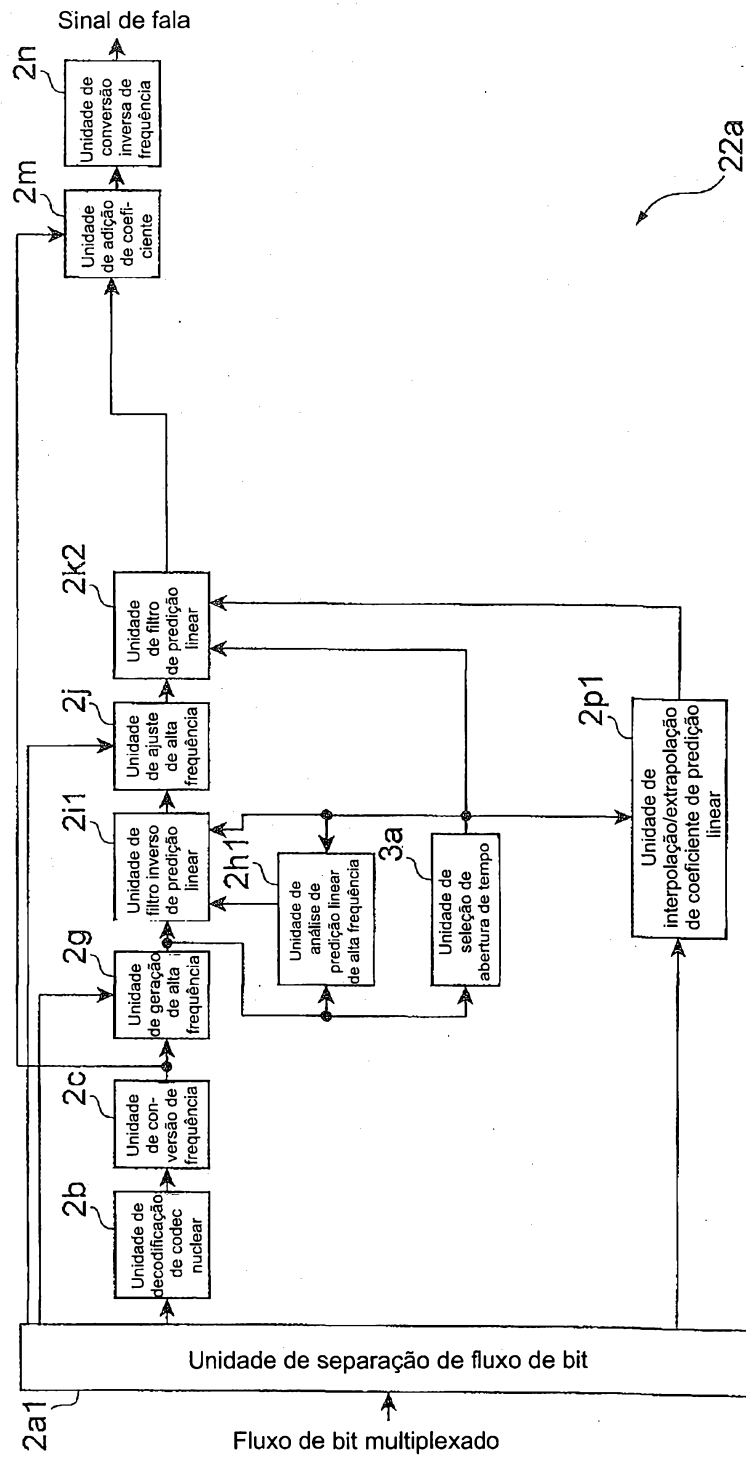


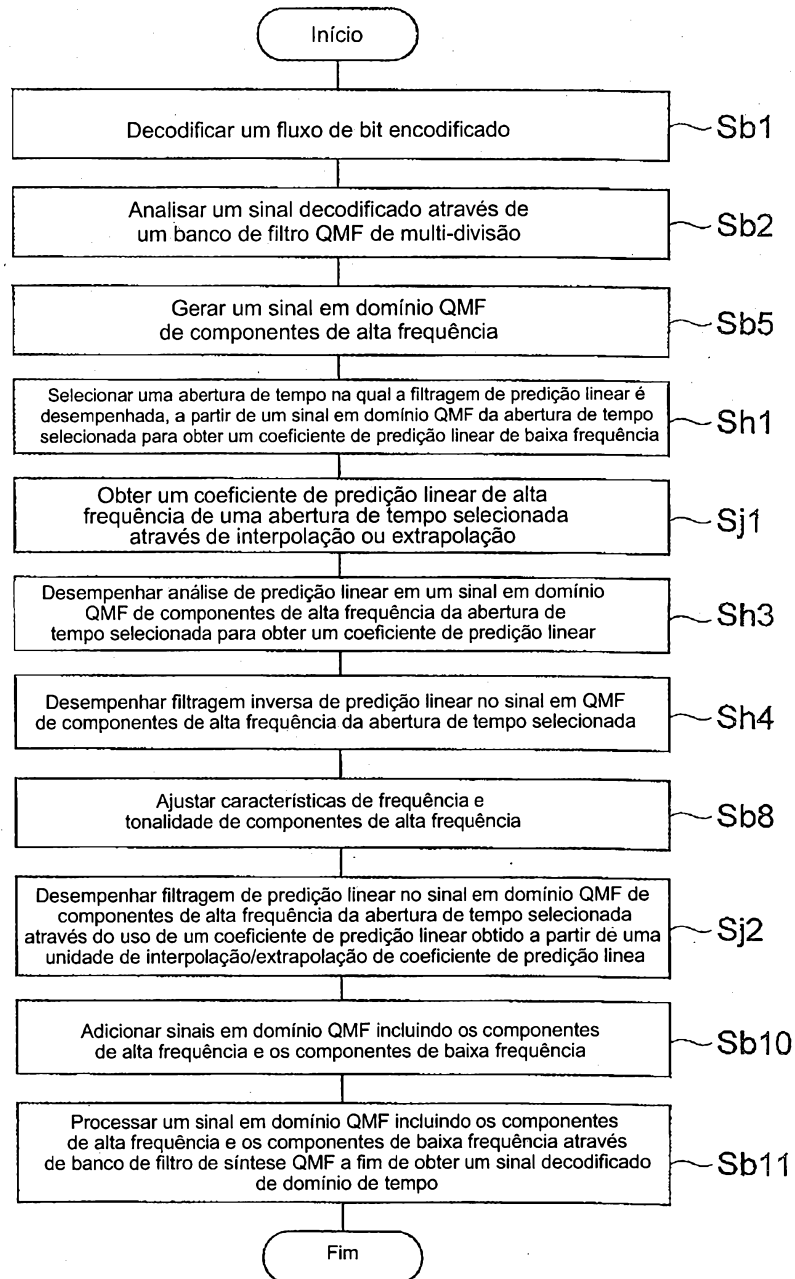
Fig.23

Fig. 24

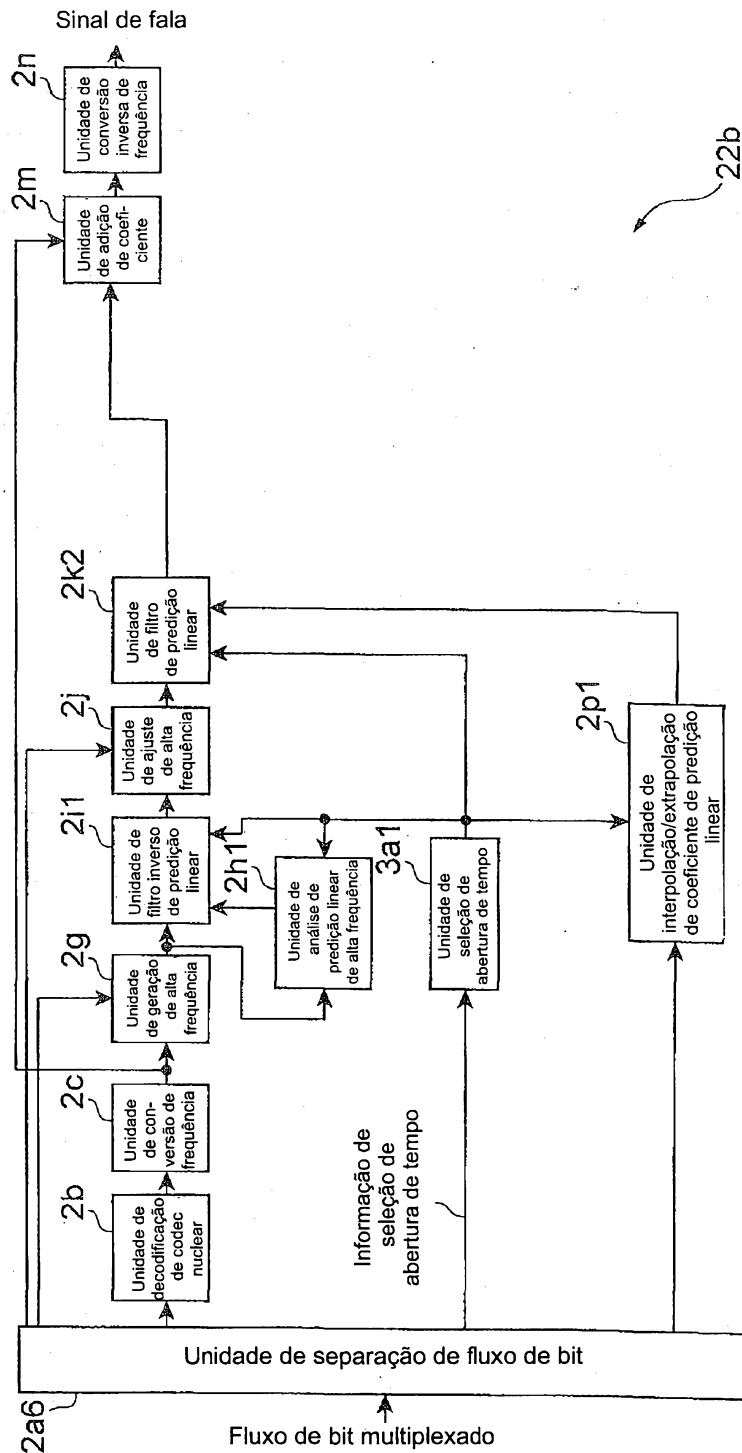


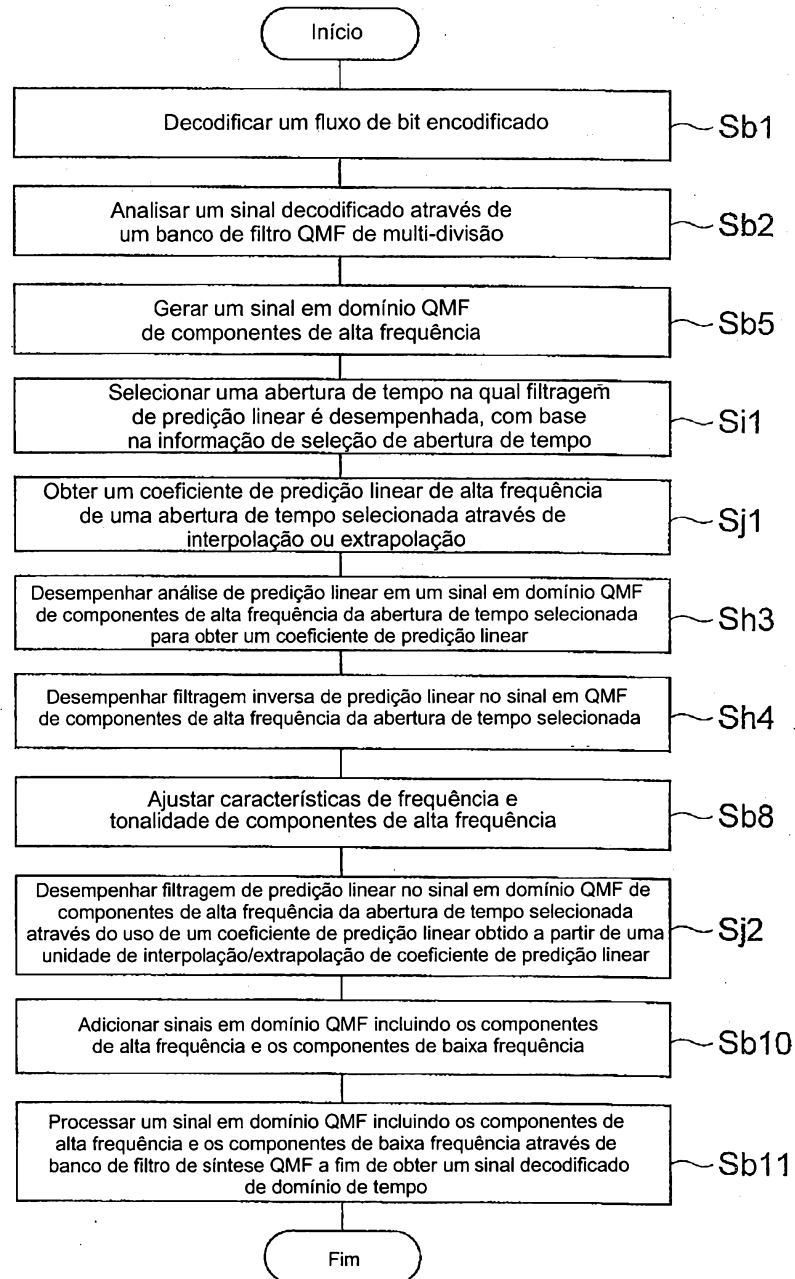
Fig.25

Fig. 26

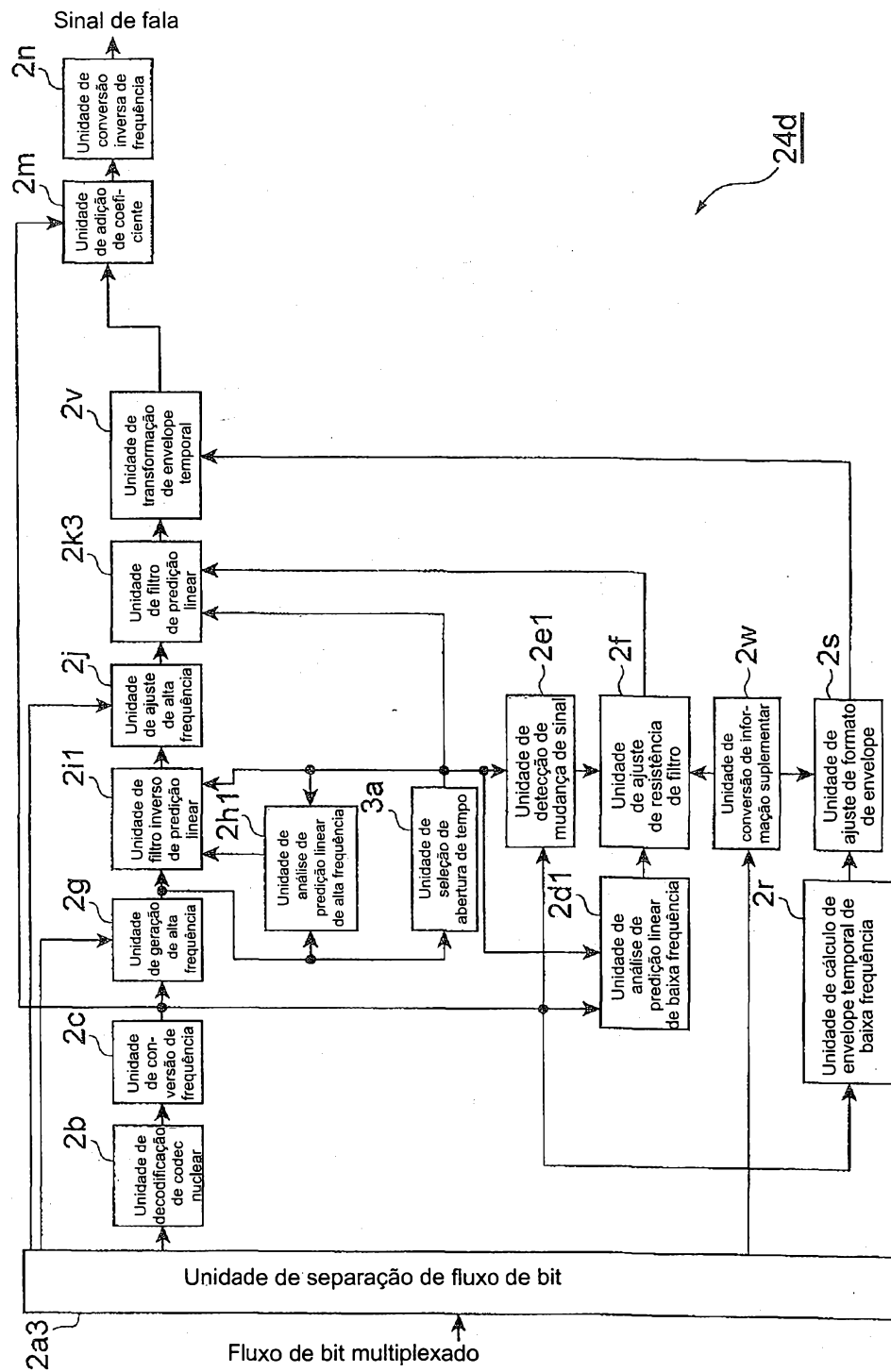


Fig.27

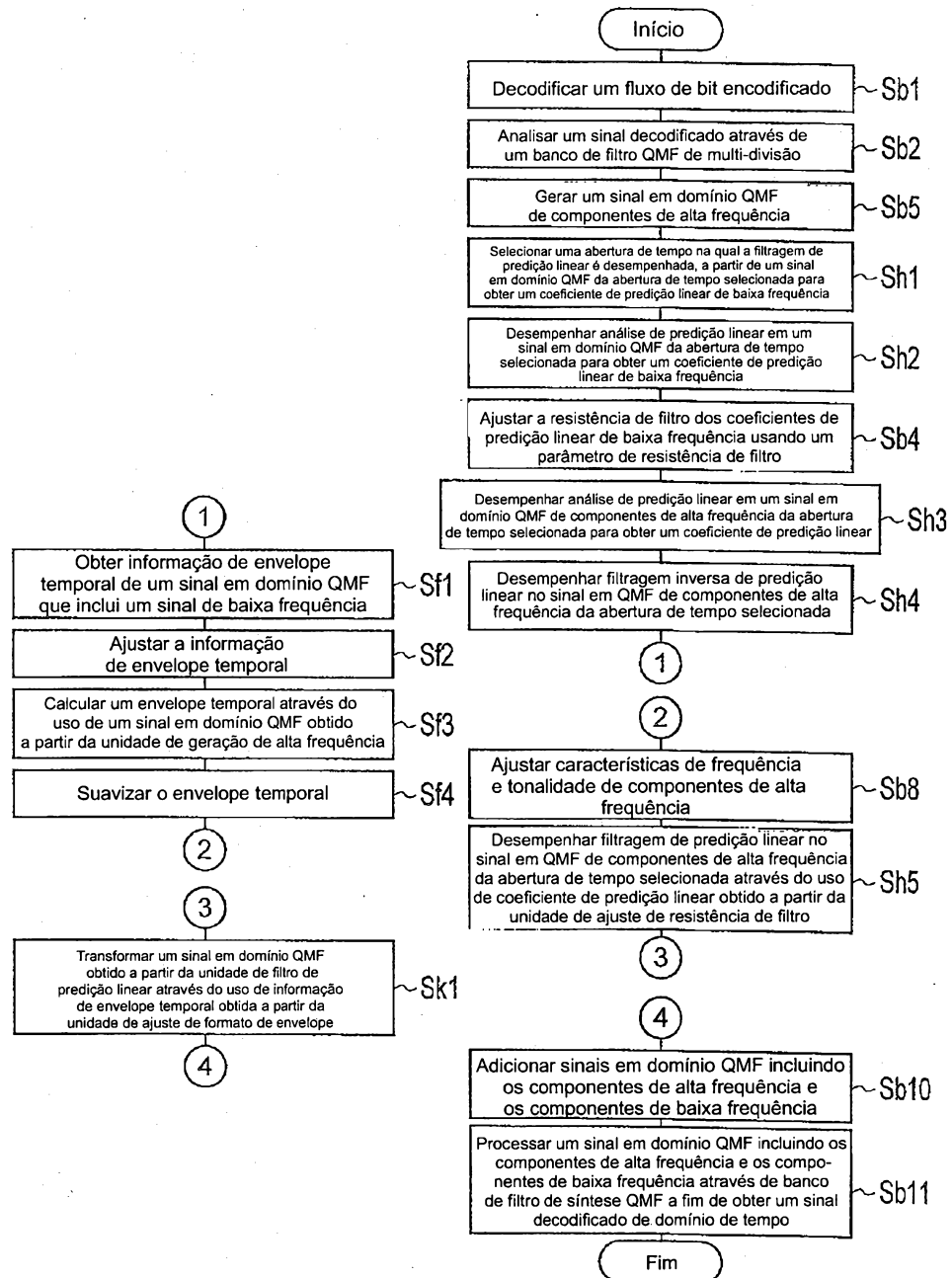


Fig. 28

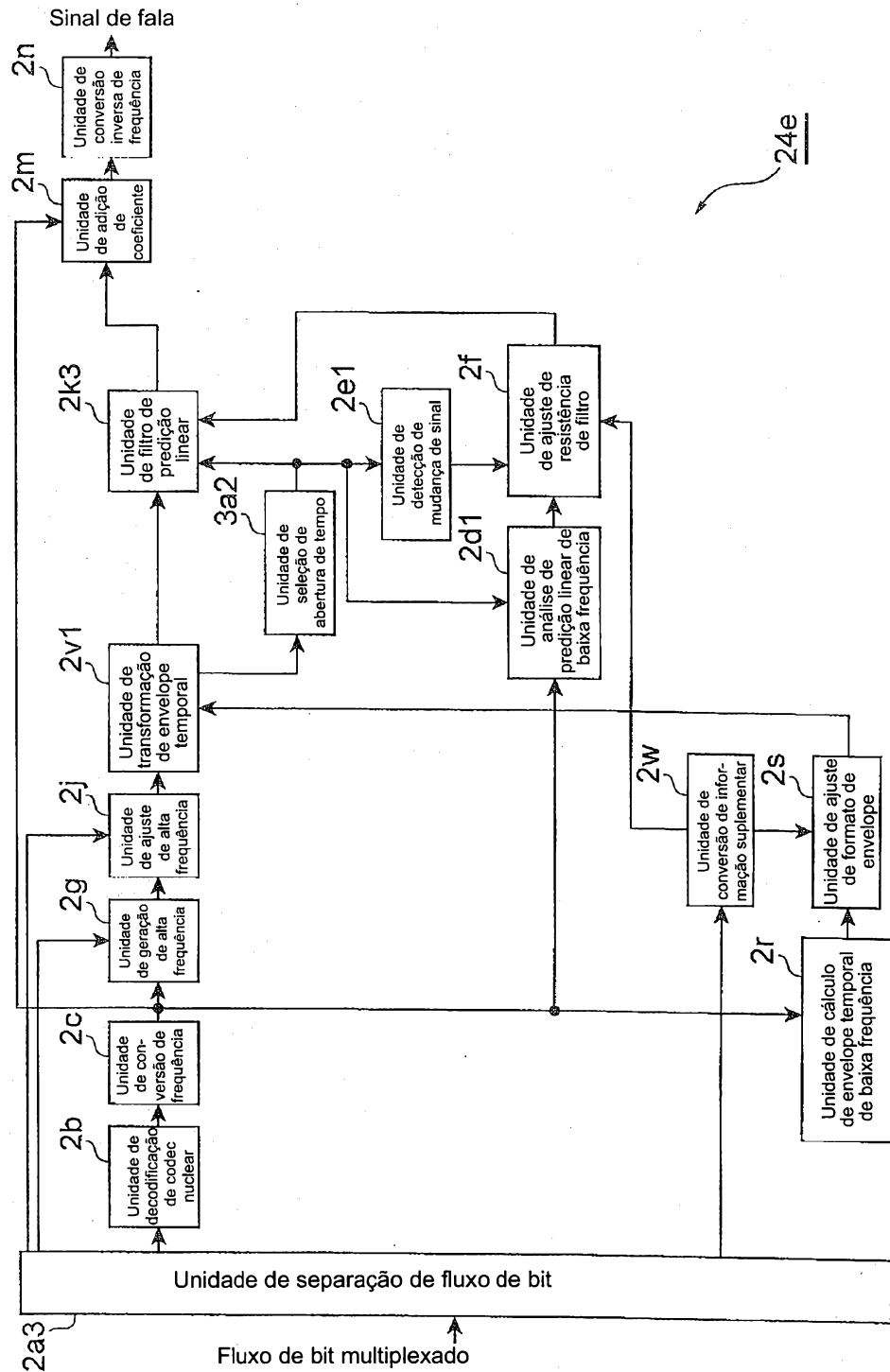


Fig.29

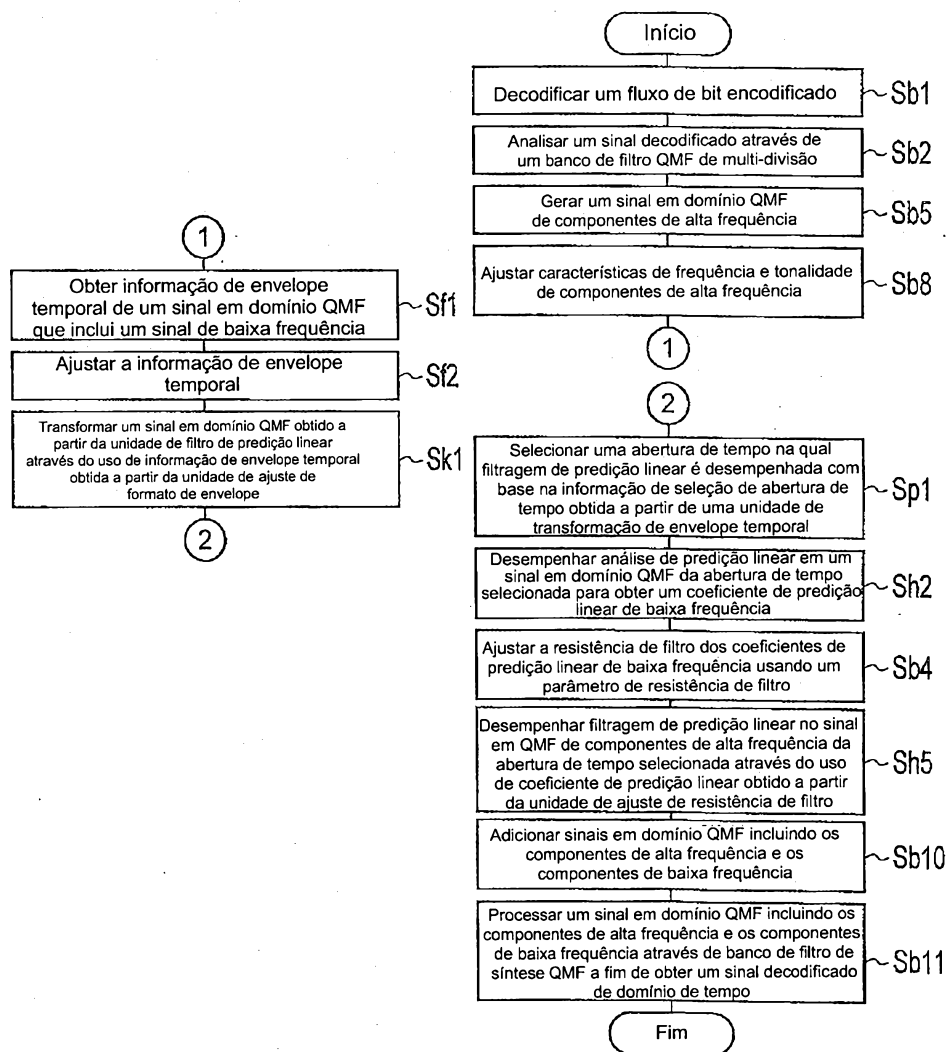
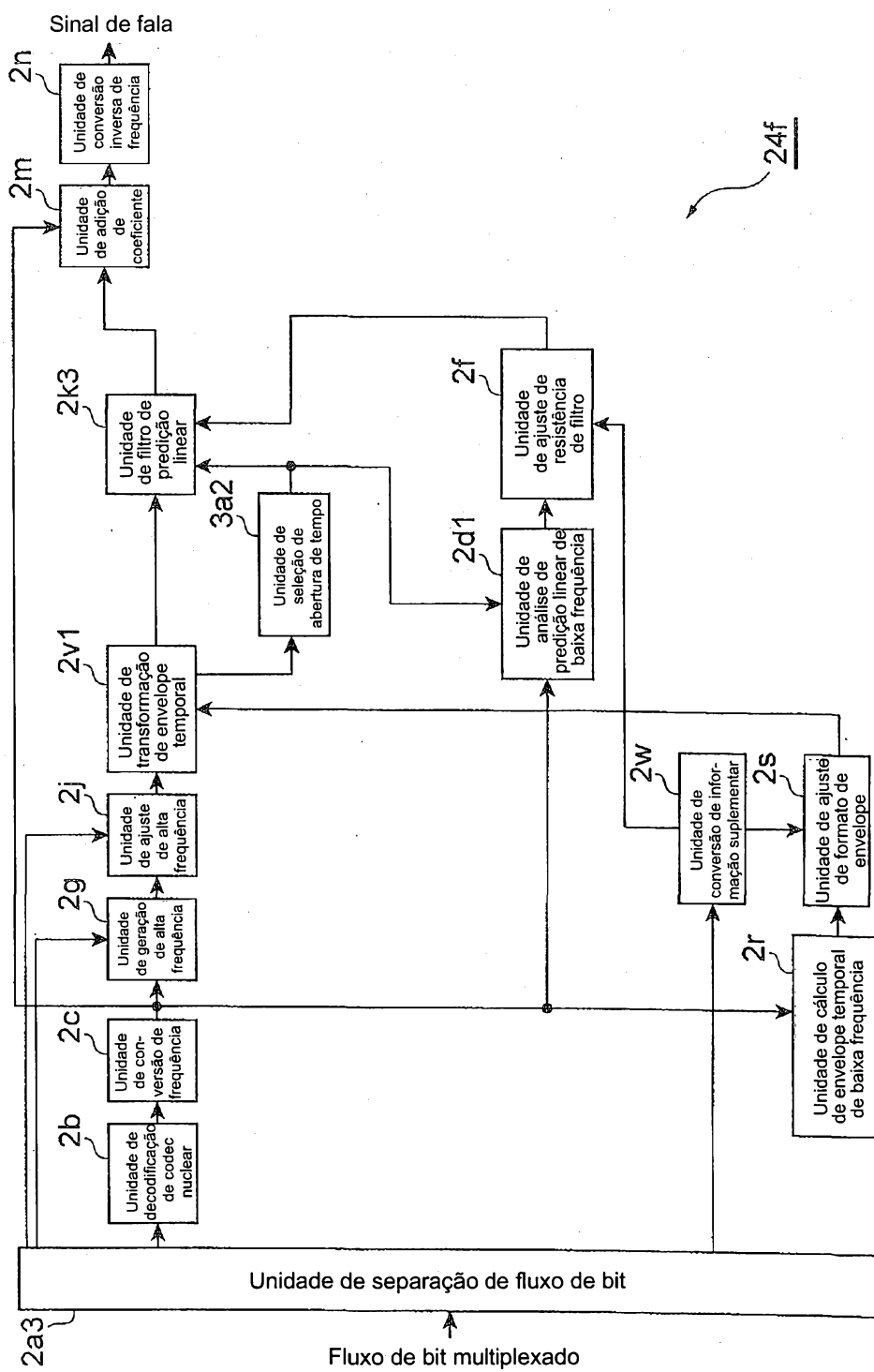


Fig.30



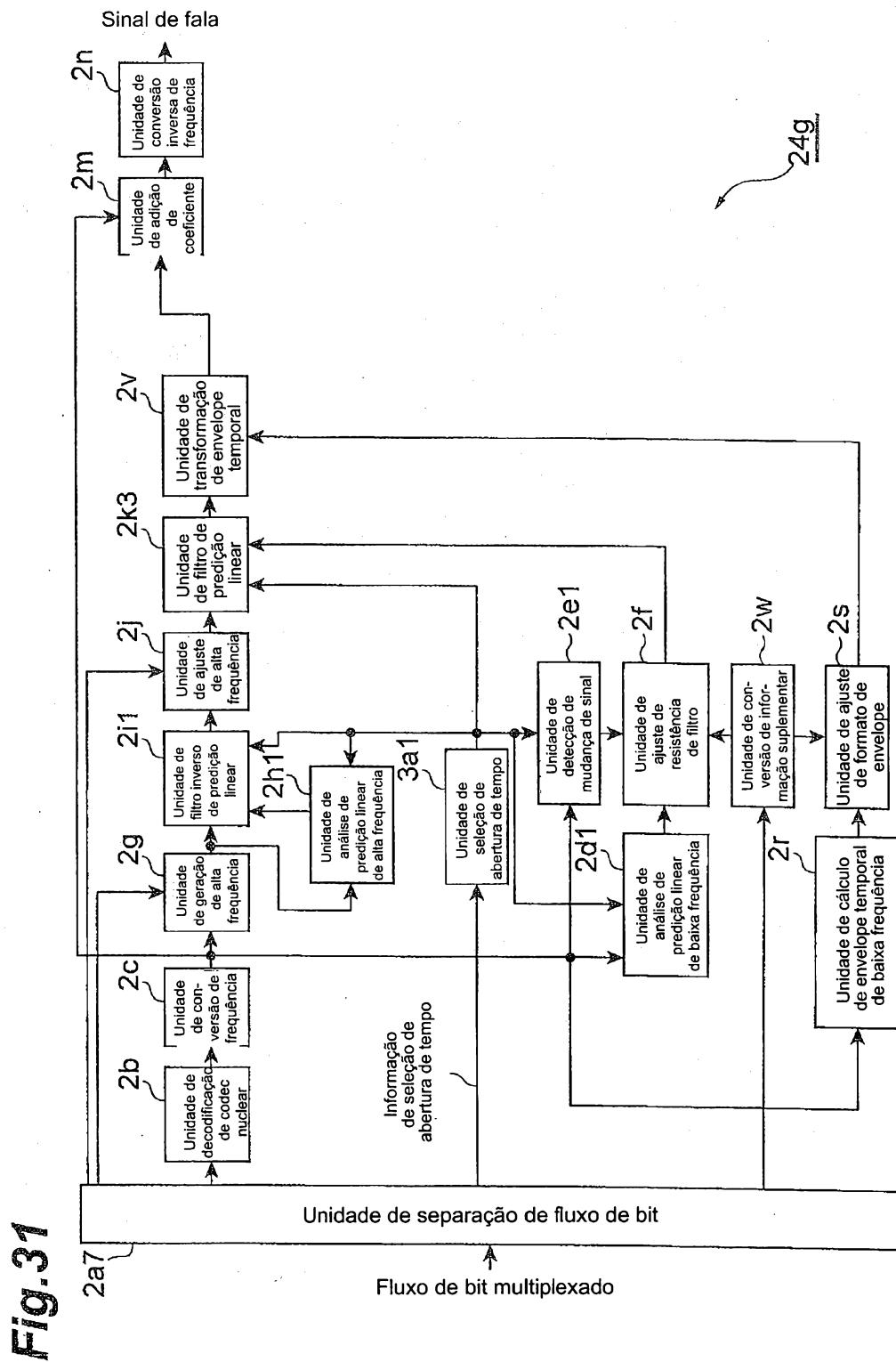


Fig.32

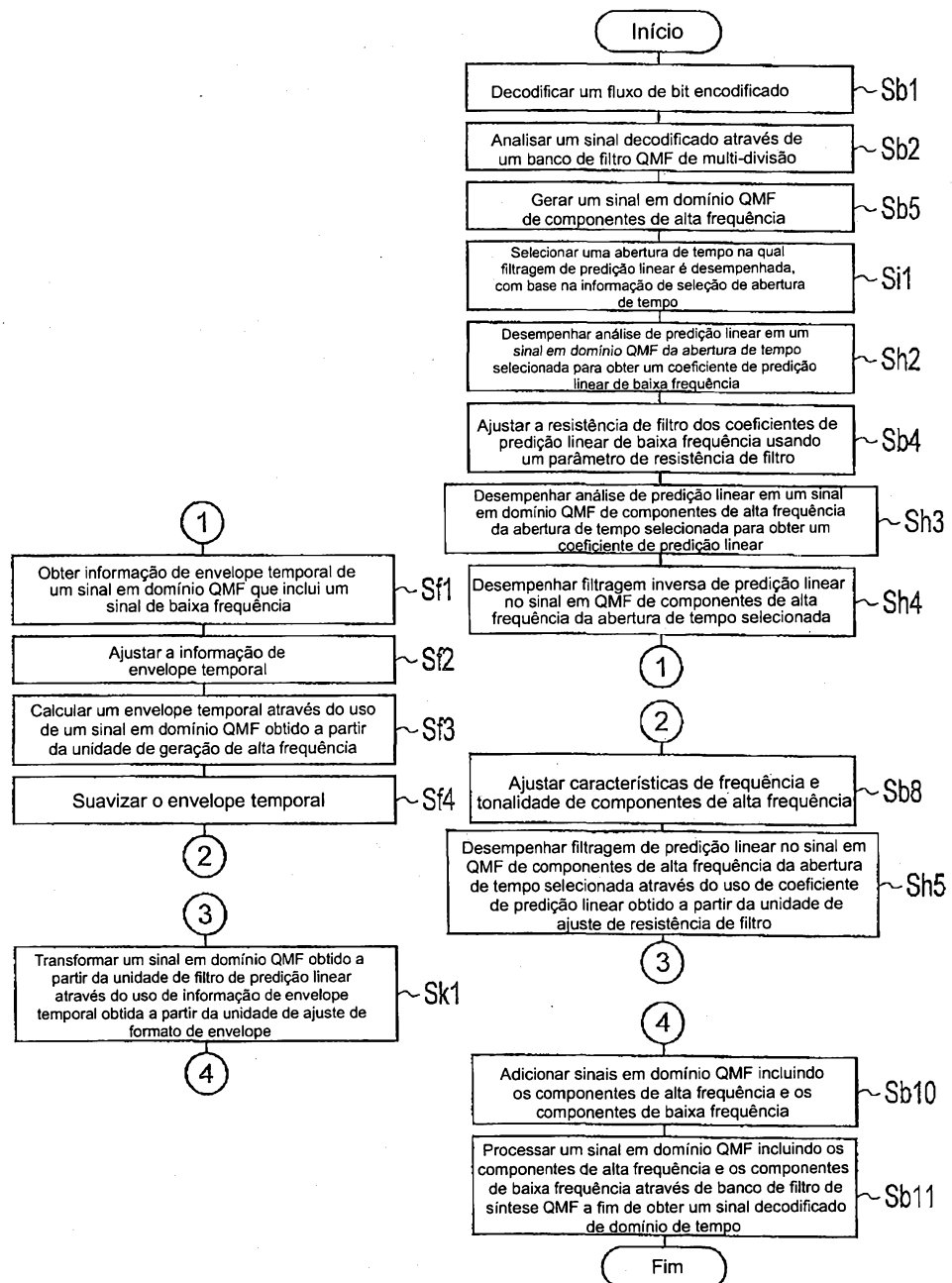


Fig.33

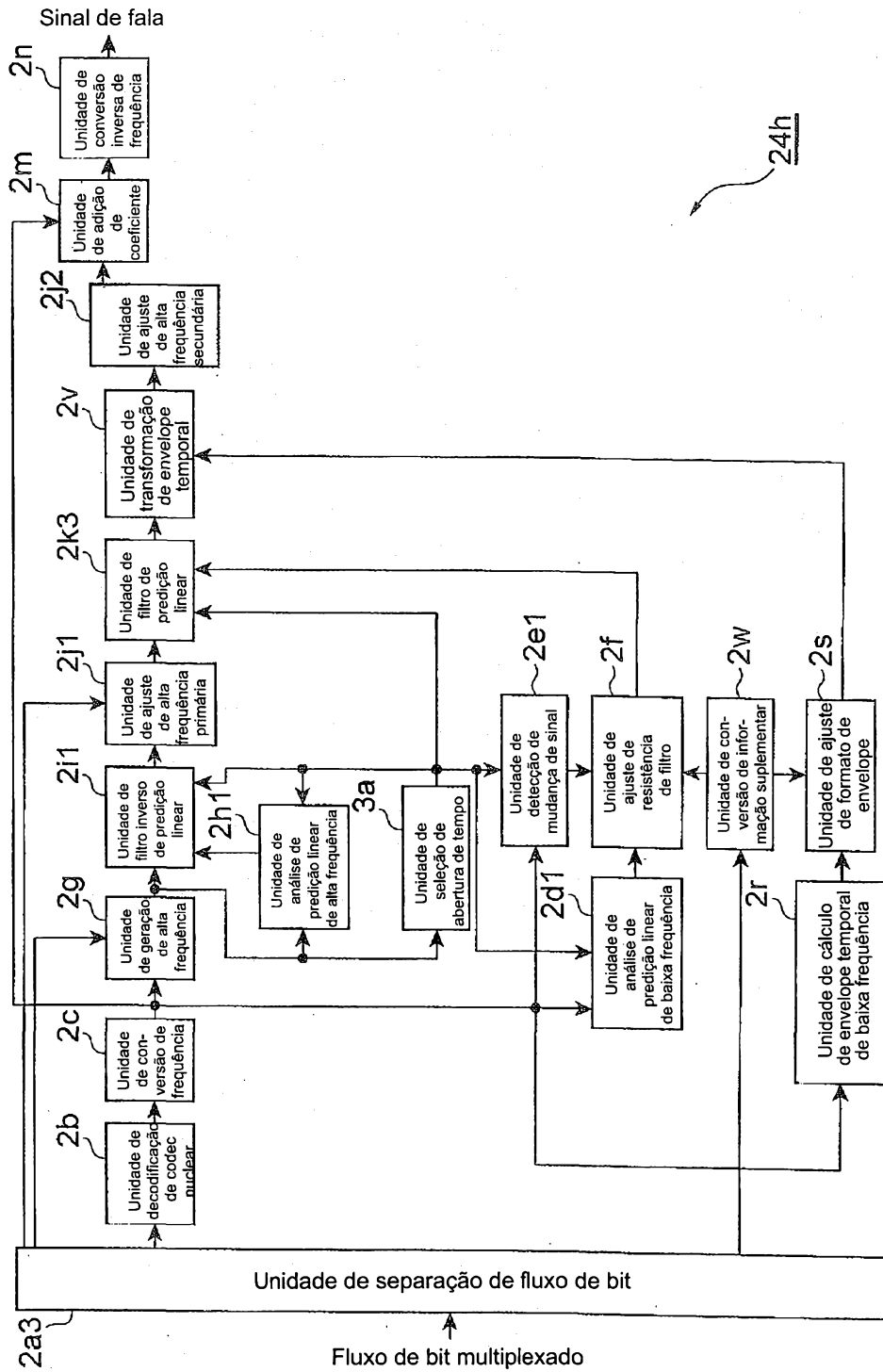


Fig.34

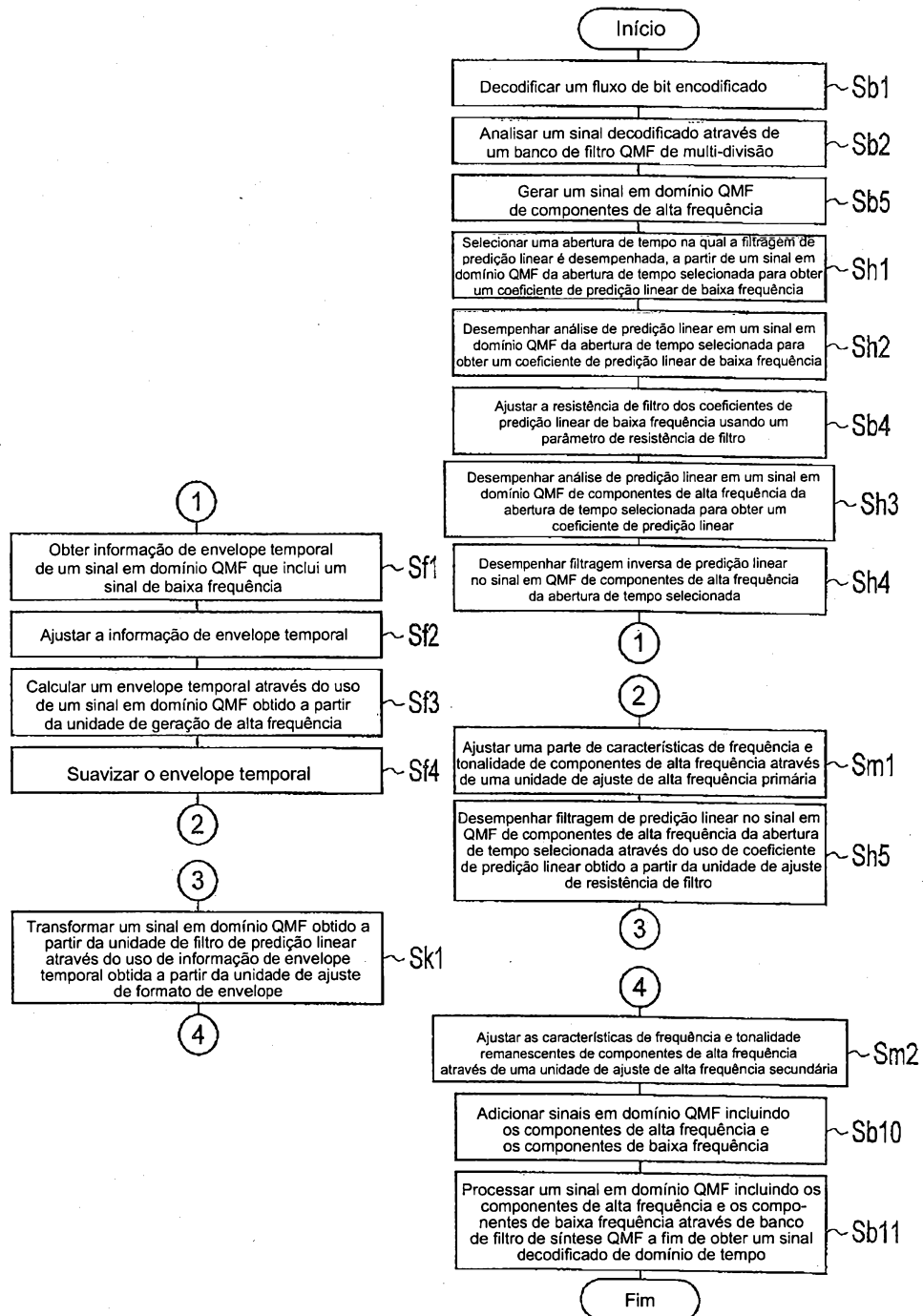


Fig.35

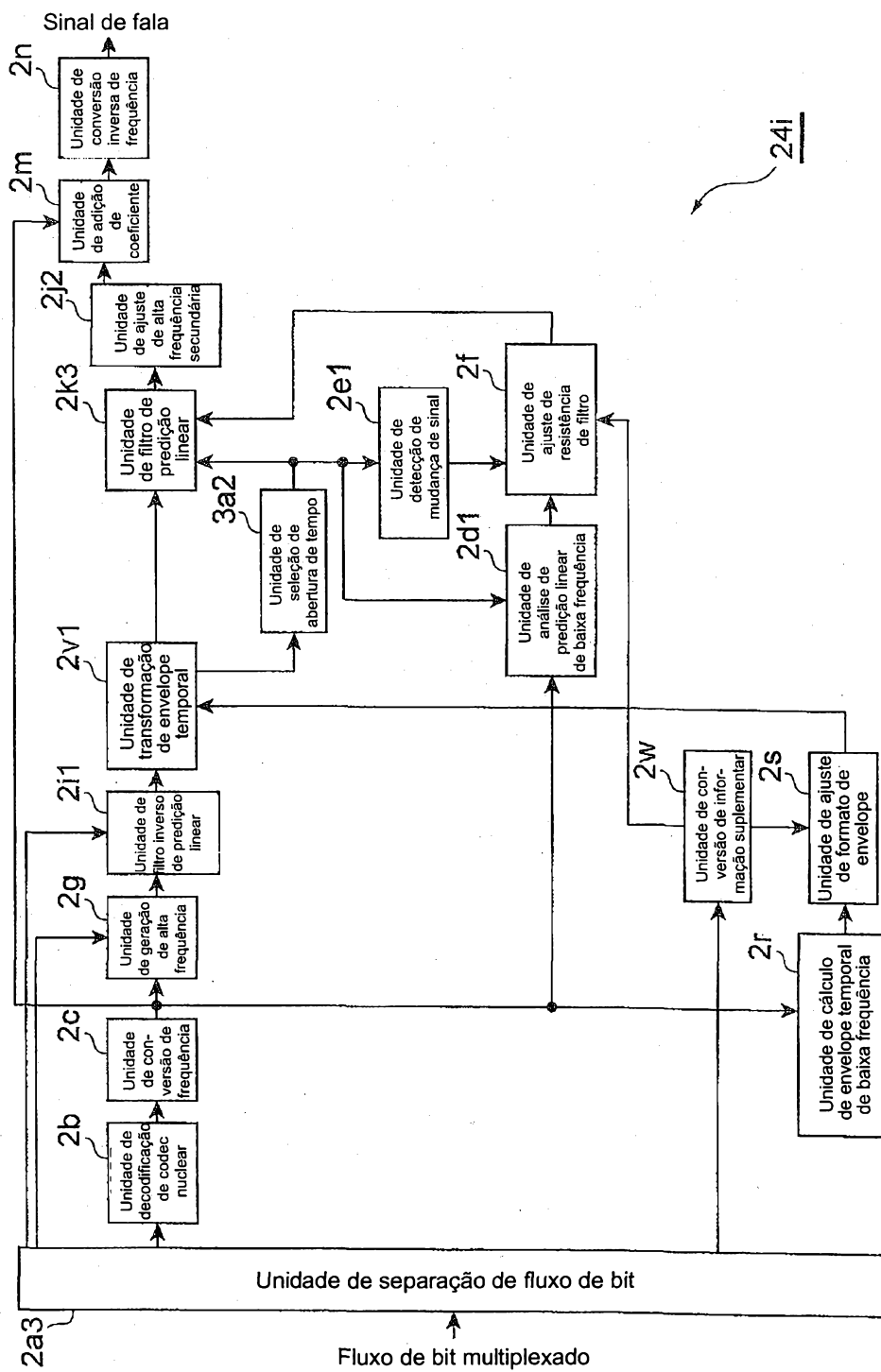
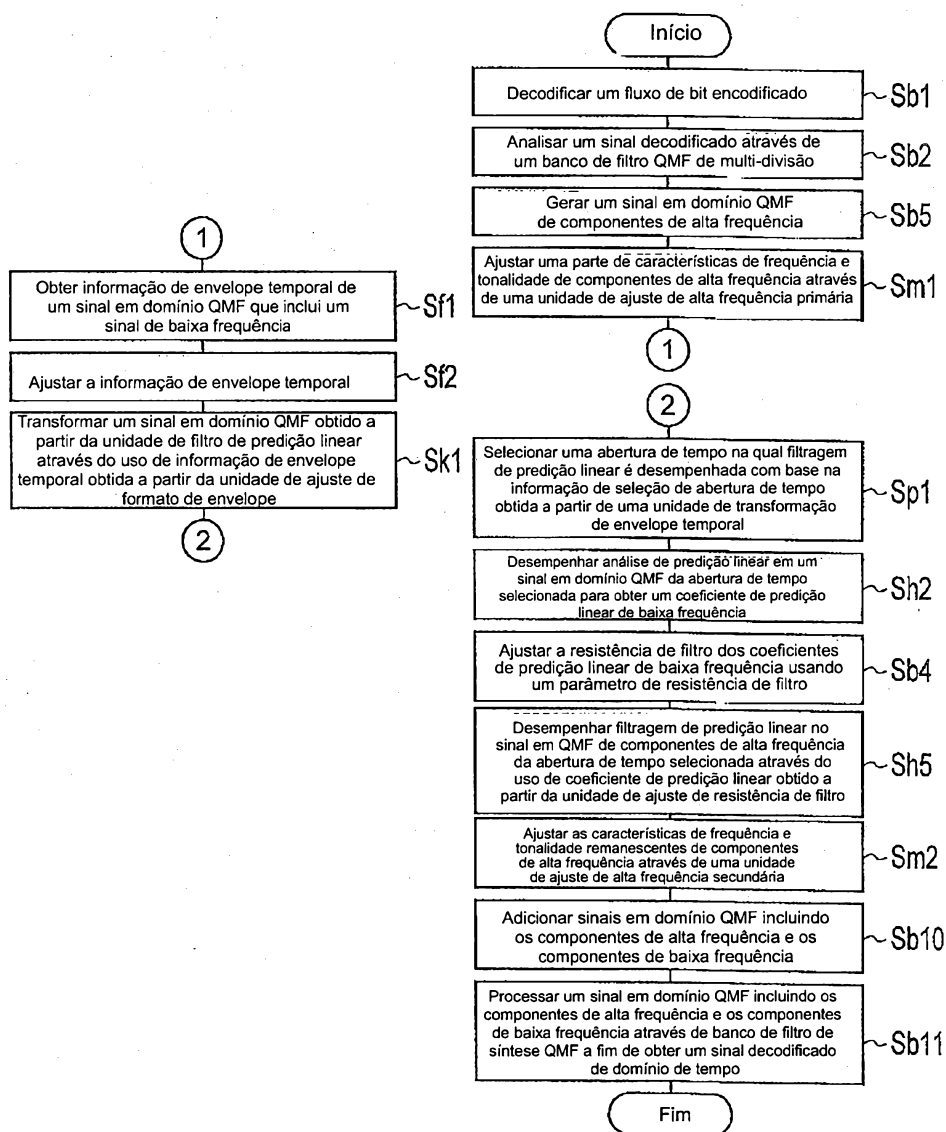
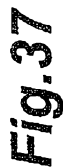


Fig.36





8301

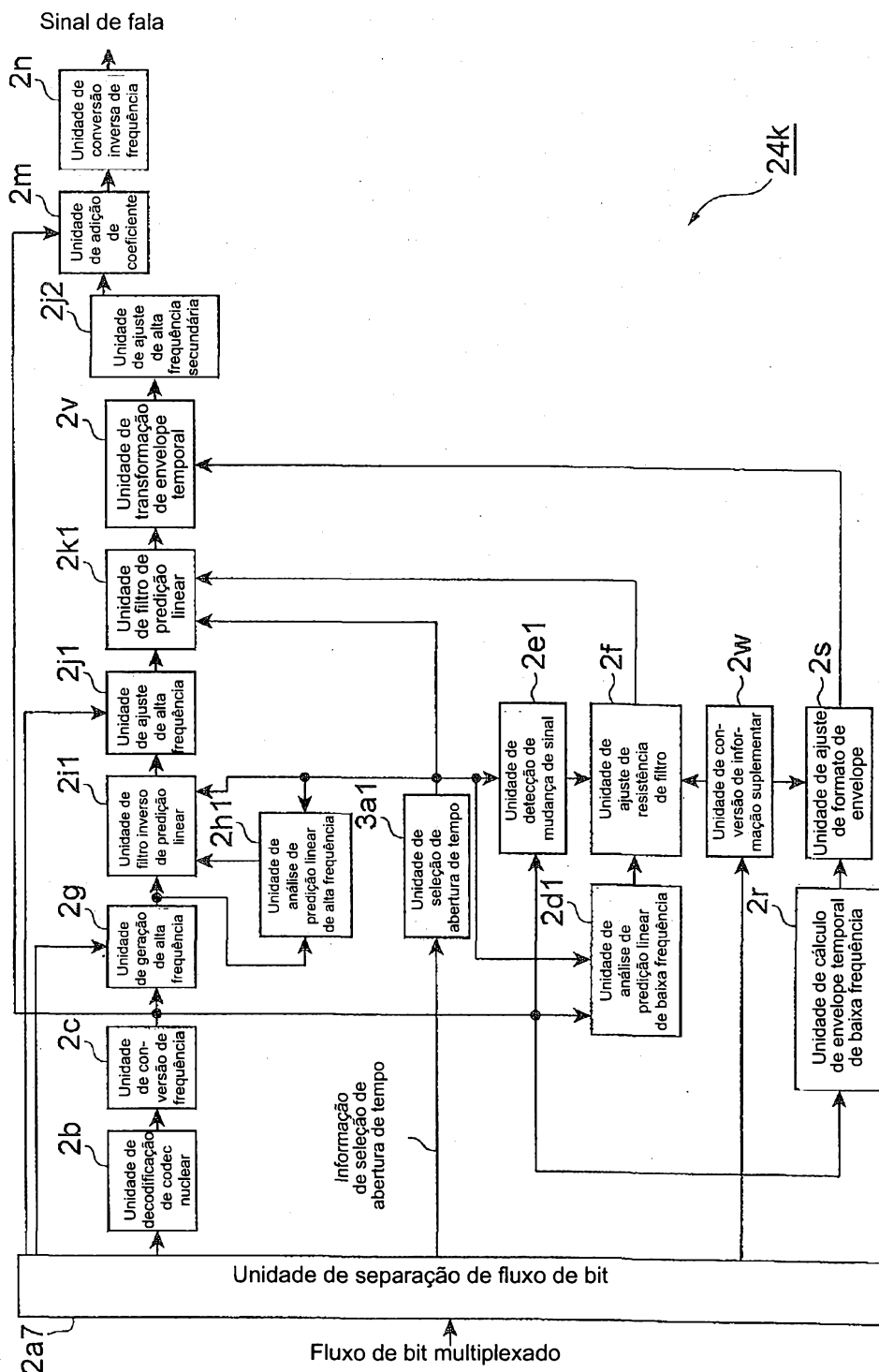


Fig.39

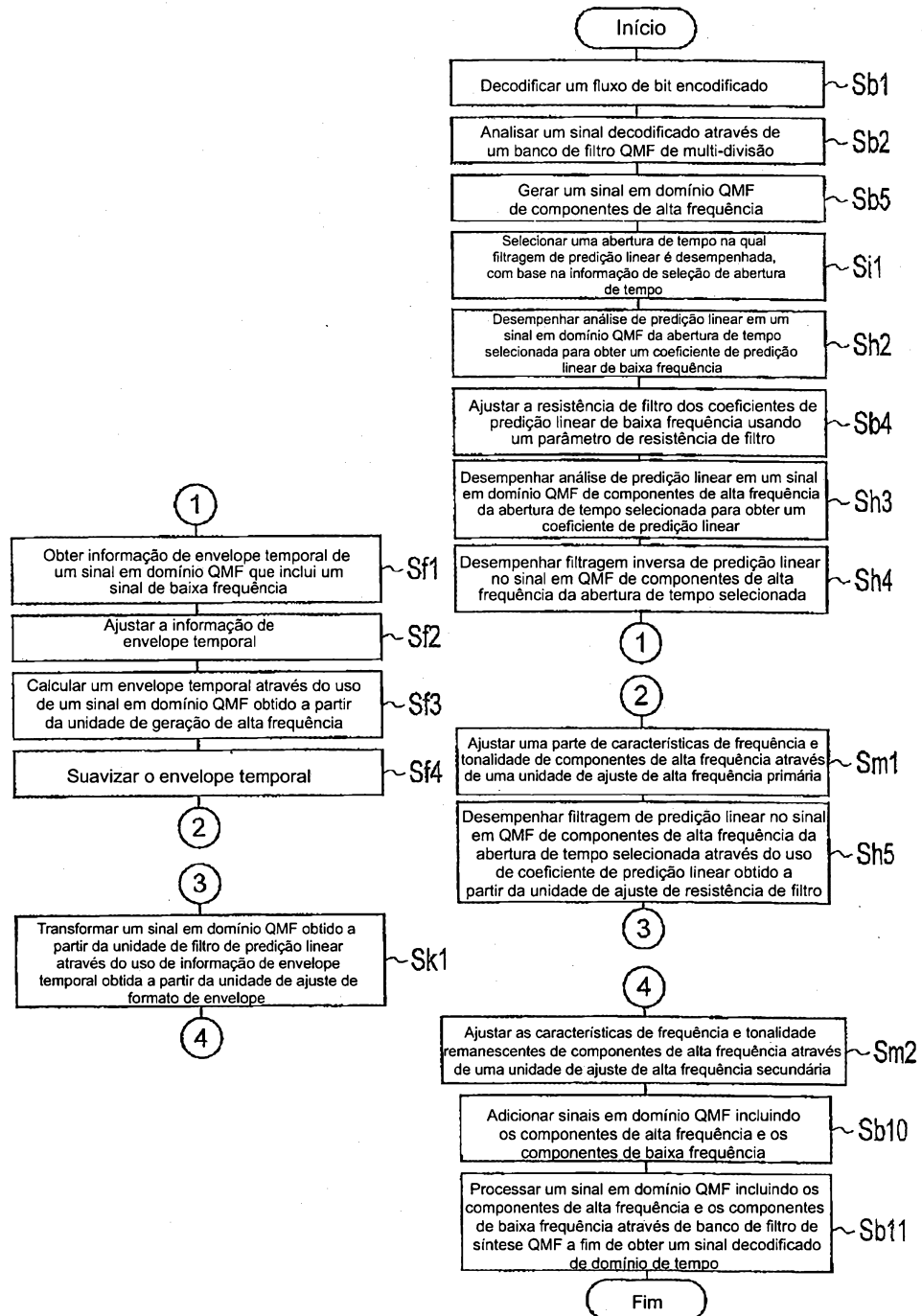


Fig.40

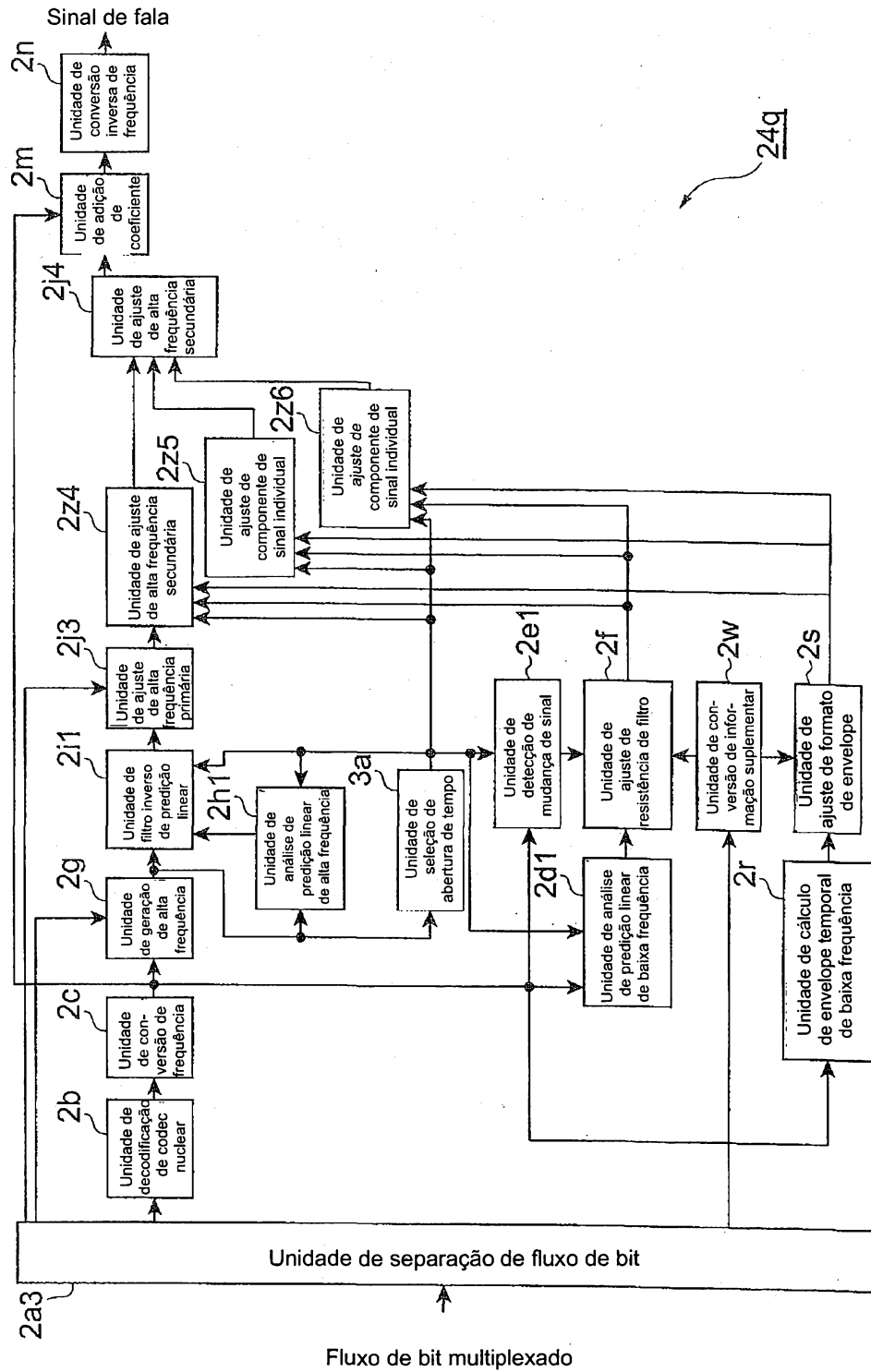


Fig.41

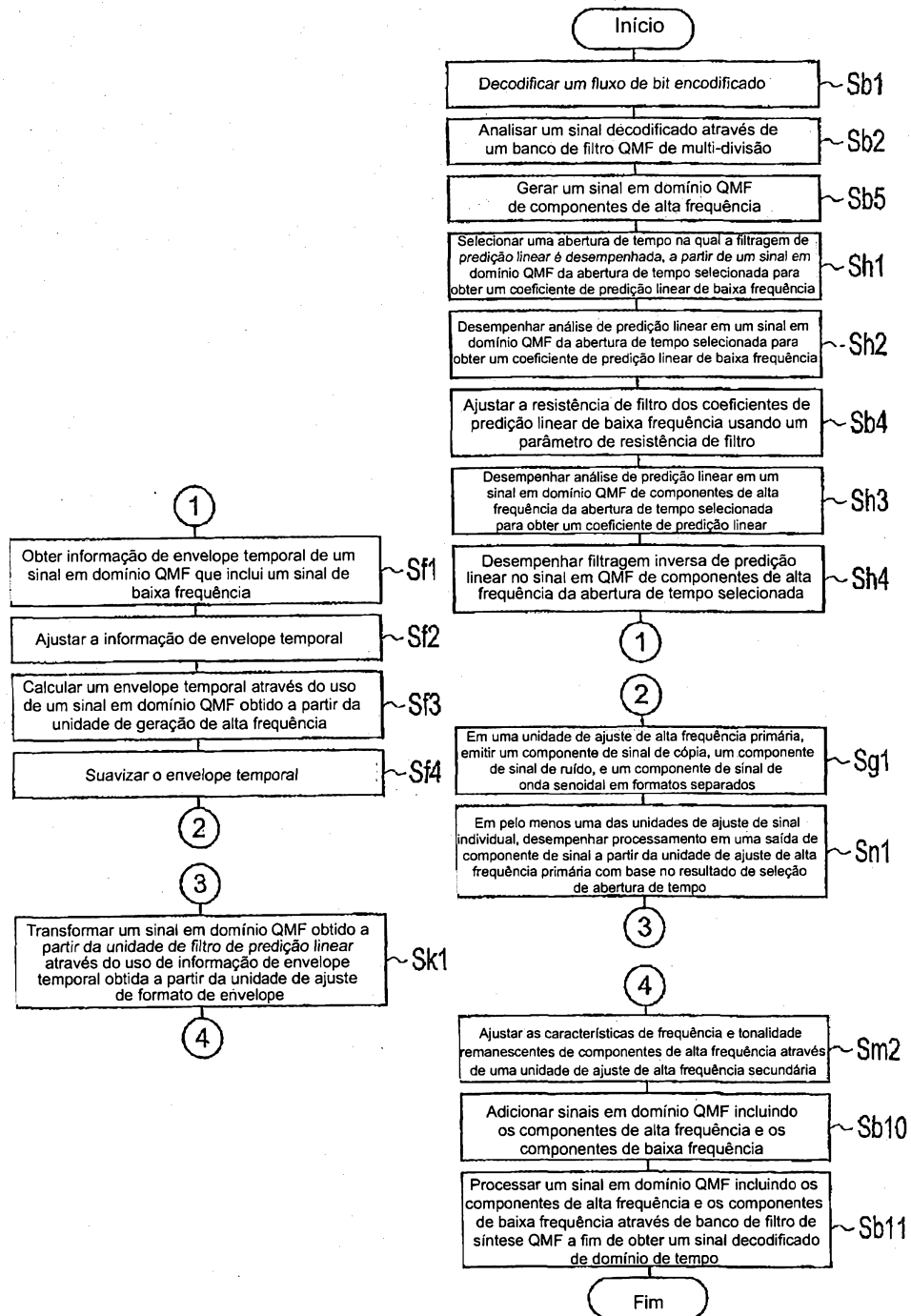


Fig.43

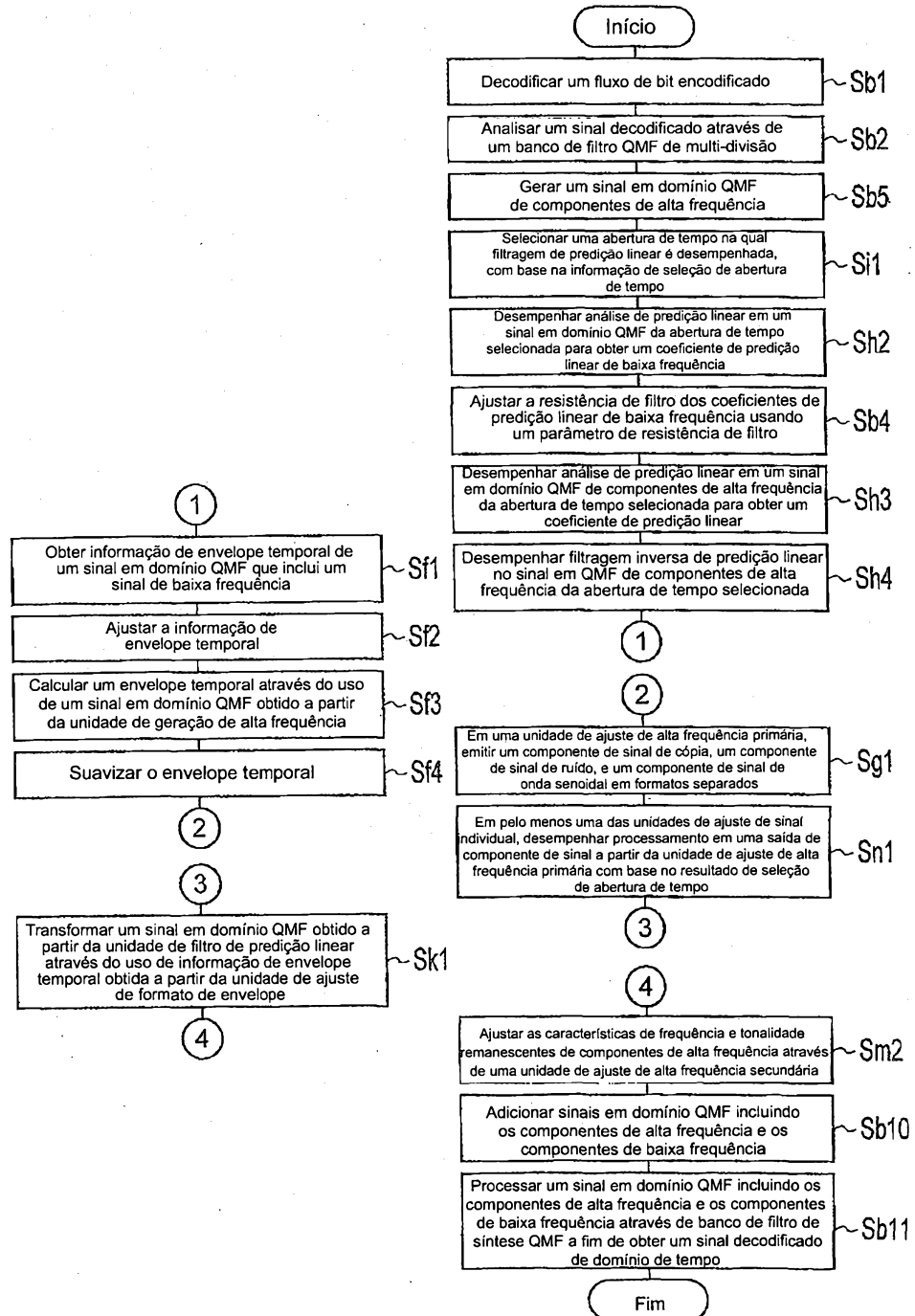


Fig. 44

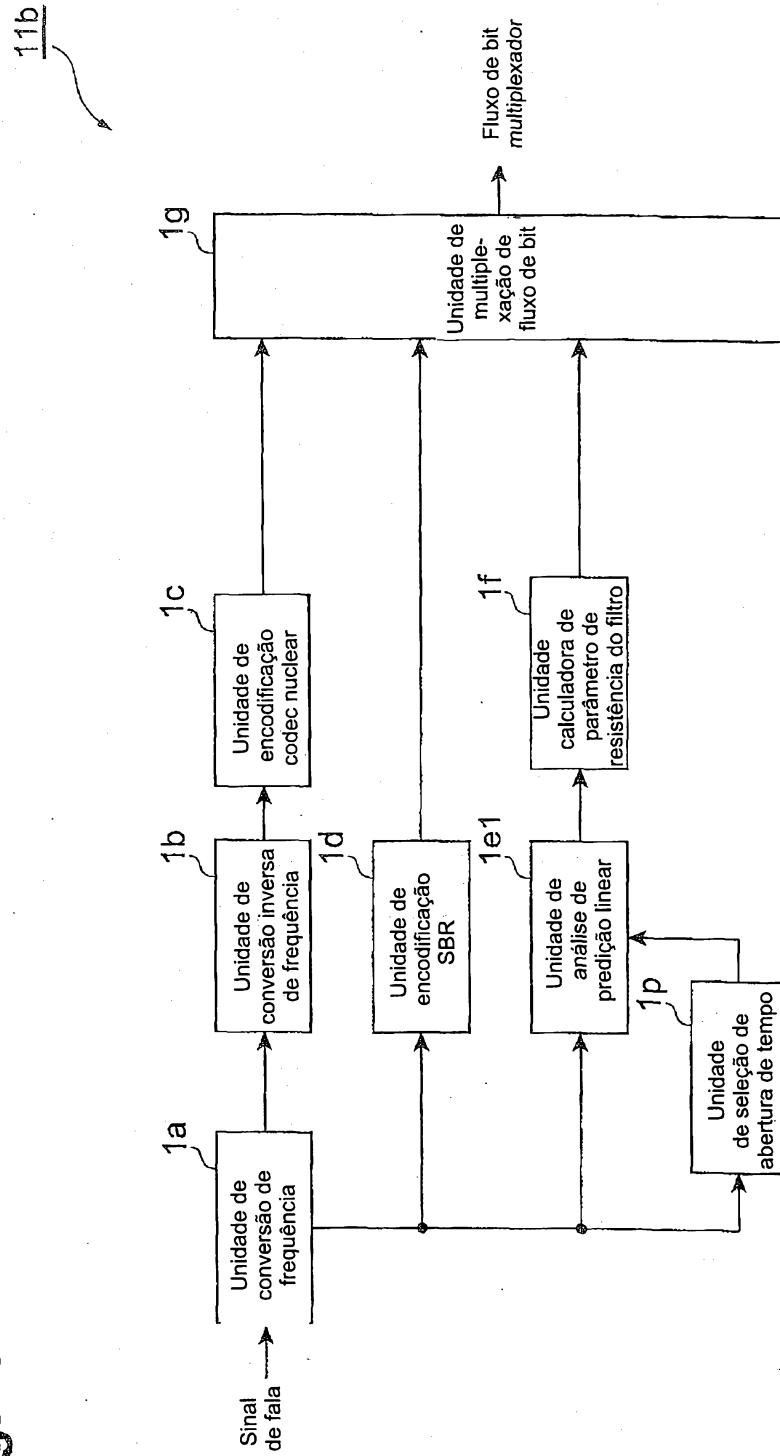


Fig.45

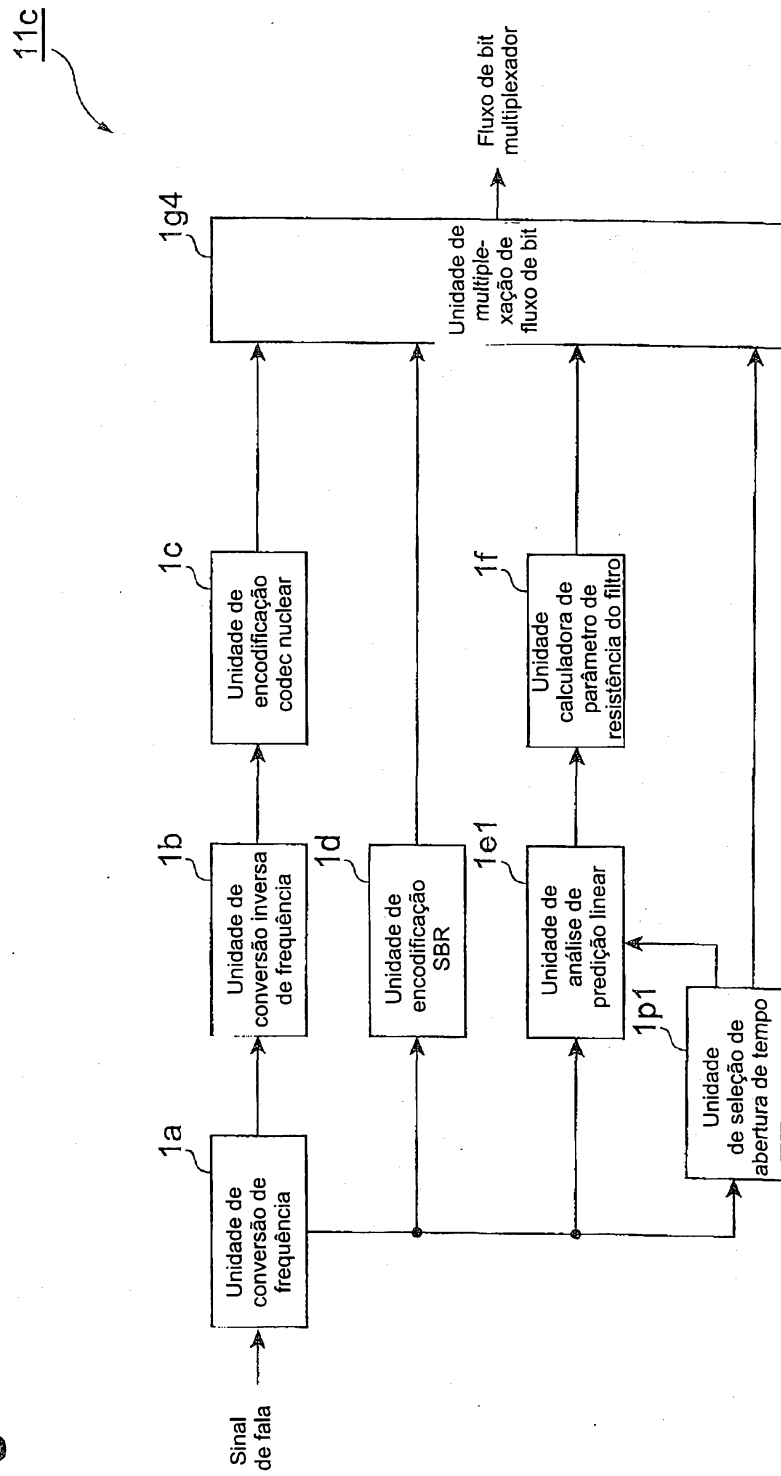


Fig.46

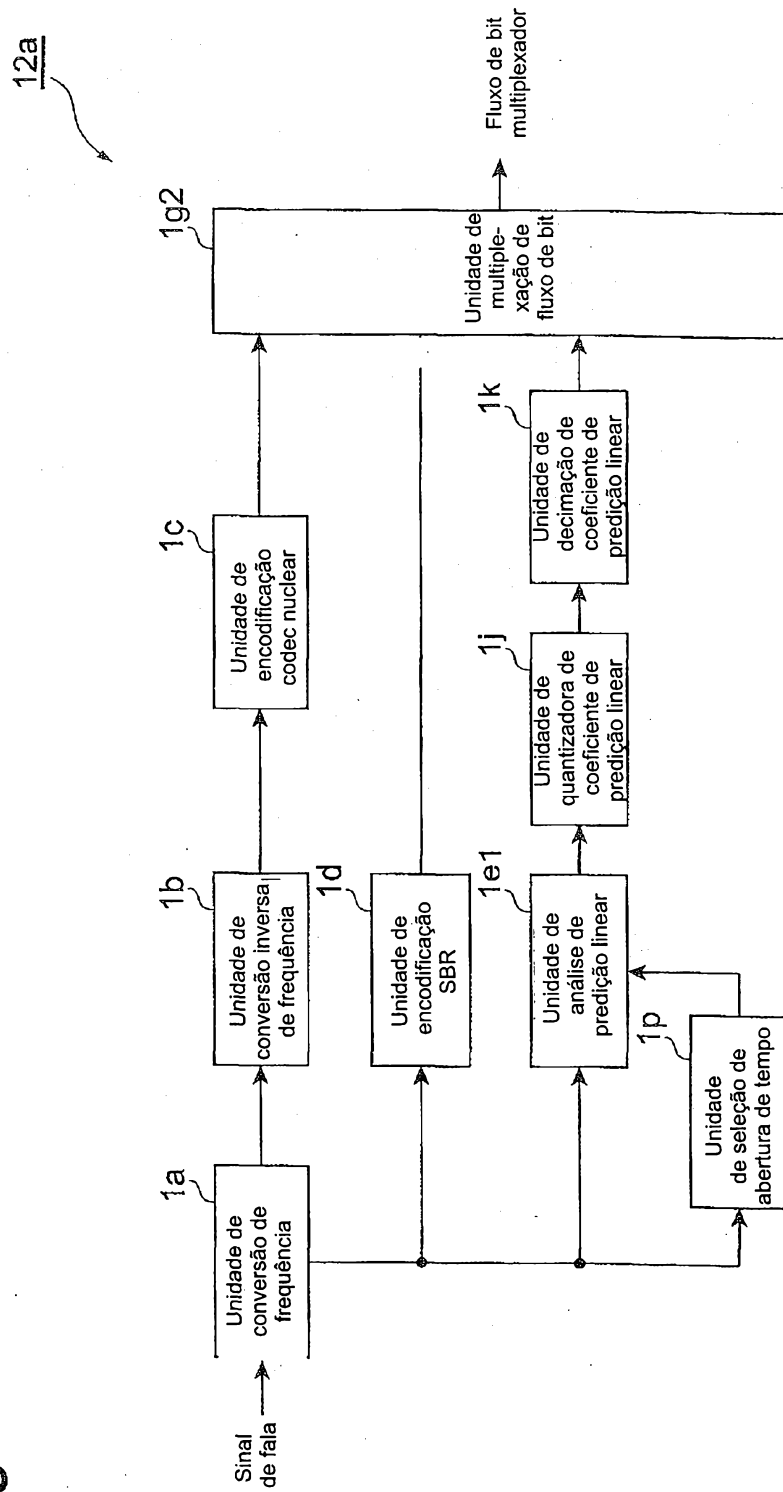


Fig. 47

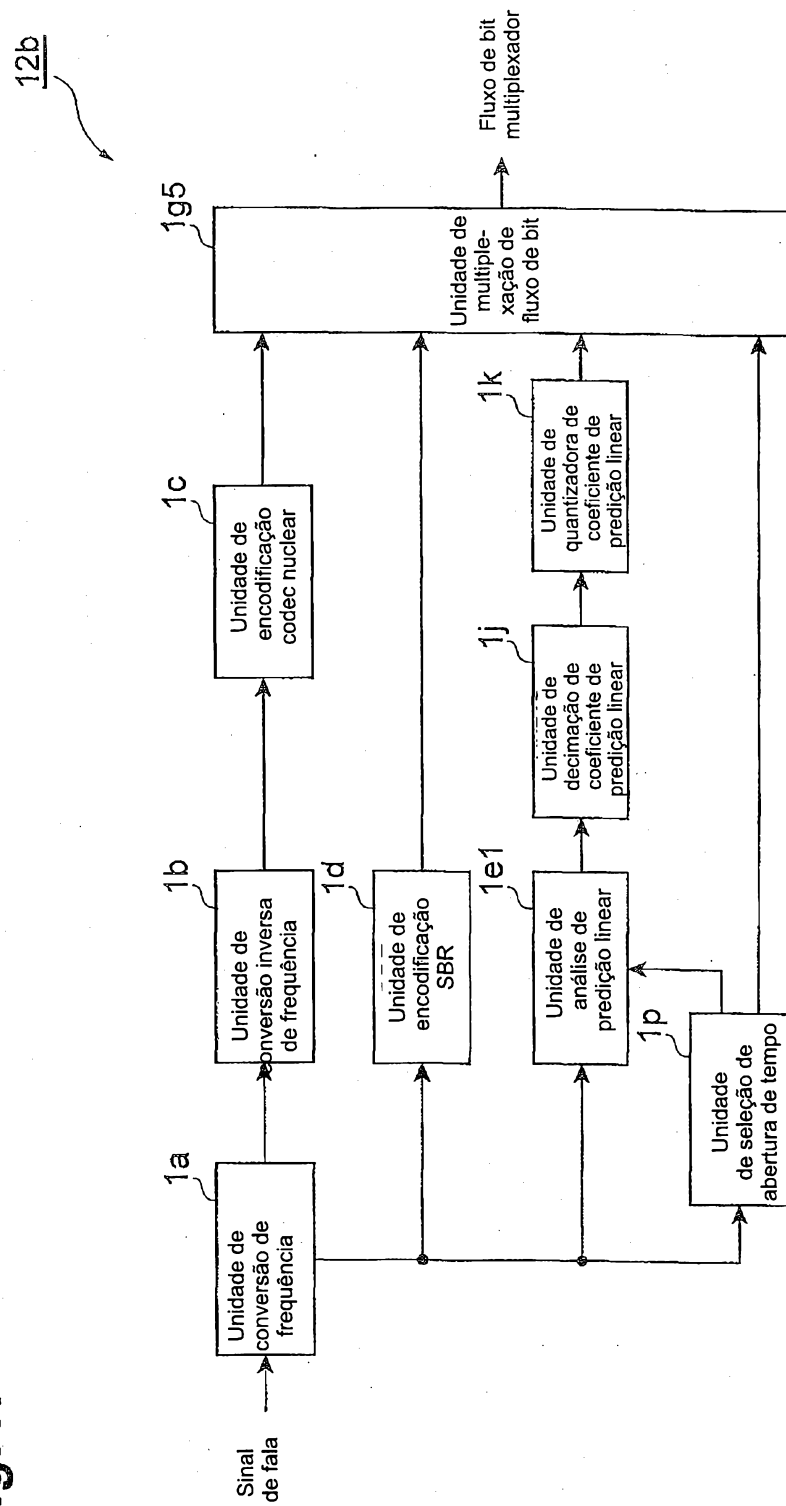


Fig. 48

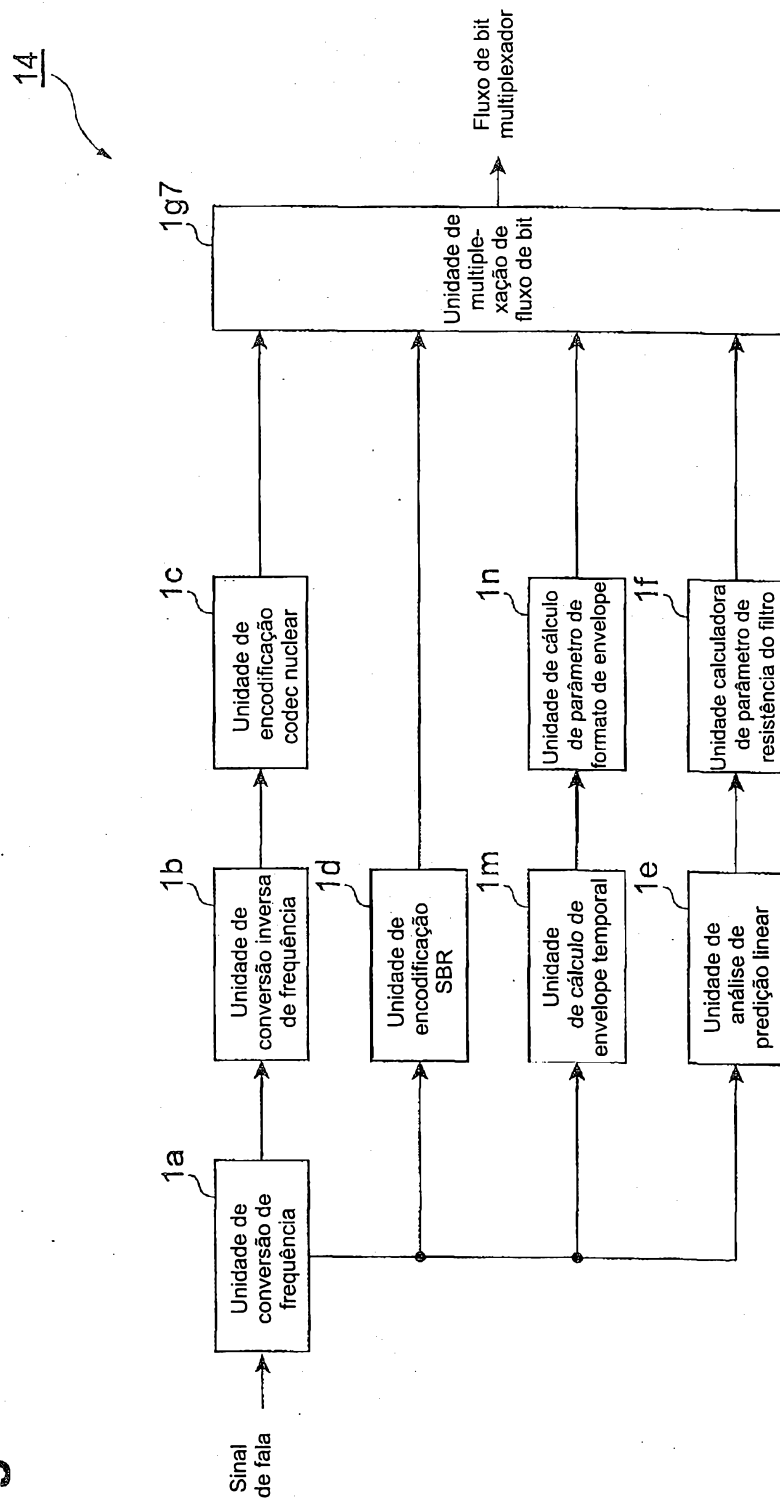


Fig. 49

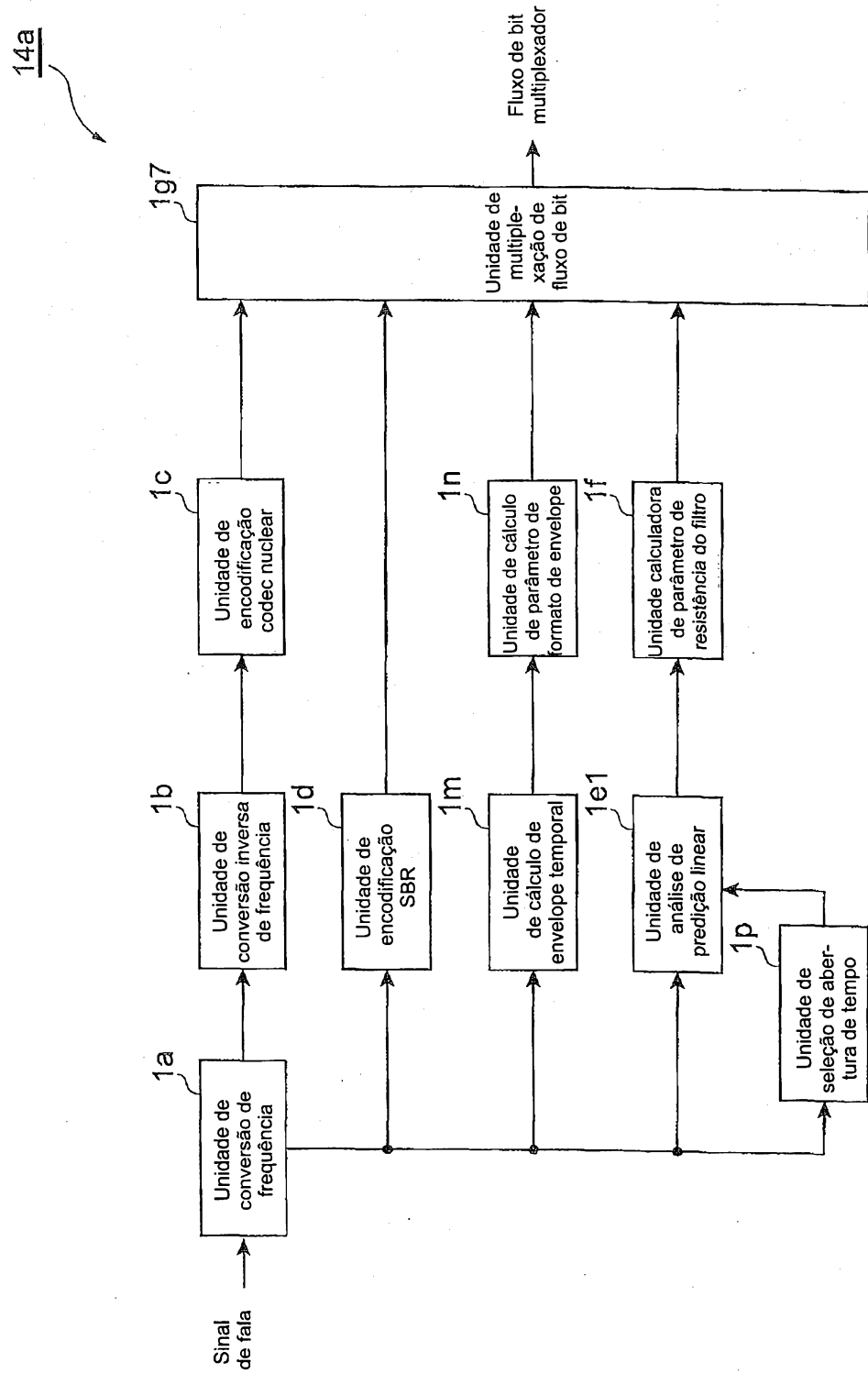
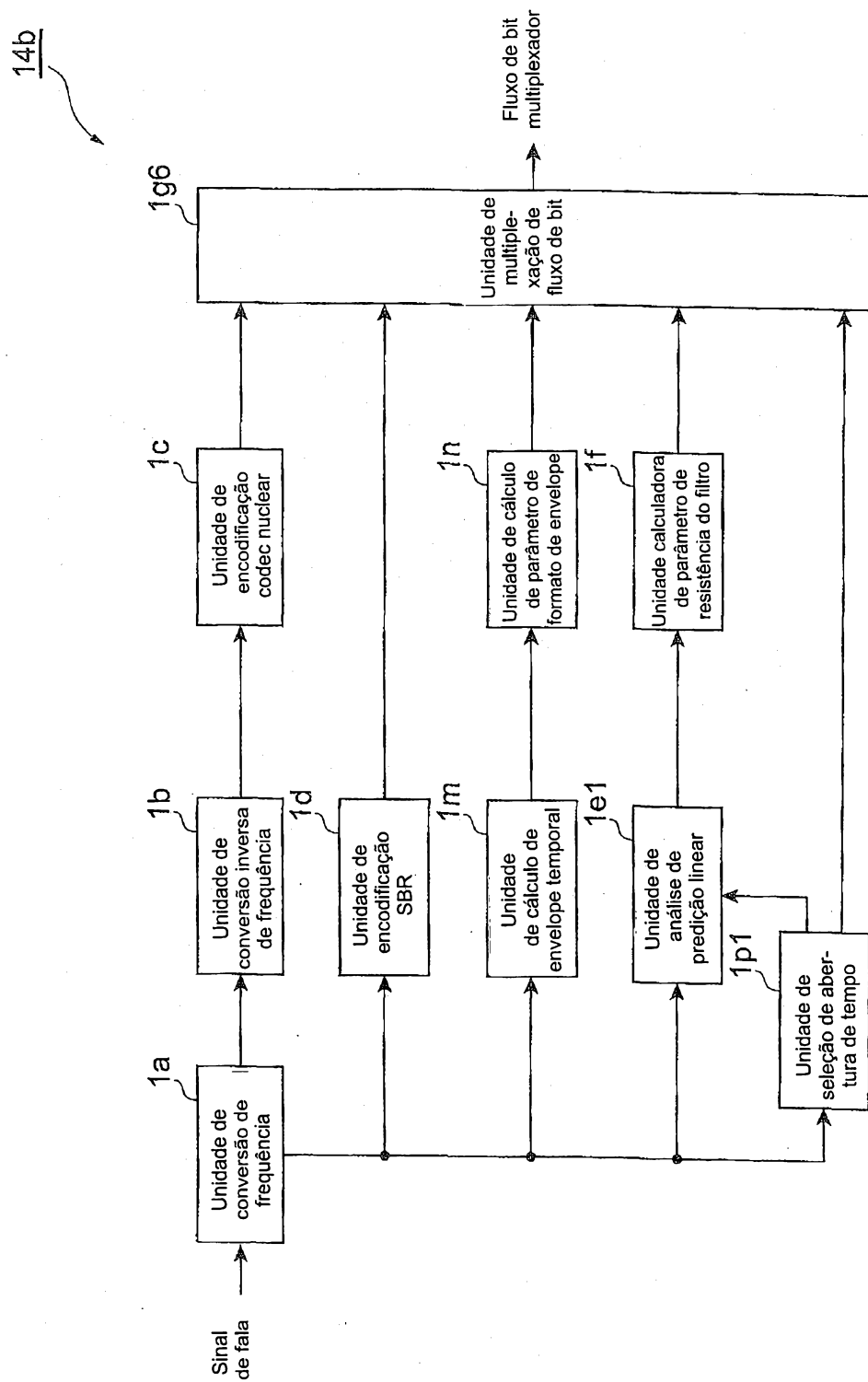


Fig. 50



RESUMO

Patente de Invenção: "DISPOSITIVO DE CODIFICAÇÃO DE VOZ, DISPOSITIVO DE DECODIFICAÇÃO DE VOZ, MÉTODO DE CODIFICAÇÃO DE VOZ, MÉTODO DE DECODIFICAÇÃO DE VOZ, PROGRAMA DE CODIFICAÇÃO DE VOZ E PROGRAMA DE DECODIFICAÇÃO DE VOZ"..

- 5 Um coeficiente de previsão linear de um sinal representado em um domínio de frequência é representado realizando-se a análise de previsão linear in a direção de frequência com uso de um método de covariância ou um método de auto-correlação. Após a resistência de filtro do coeficiente
- 10 de previsão linear obtido ser ajustada, a filtragem é realizada na direção de frequência no sinal com uso do coeficiente ajustado, assim o envelope temporal do sinal é transformado. Isto reduz a ocorrência de pré-eco e pós-eco e melhora a qualidade subjetiva do sinal decodificado, sem aumentar significativamente a taxa de bit em uma técnica de extensão de banda no domínio
- 15 de frequência representado por SBR.