



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

(11) PI 1003636-9 BI



(22) Data do Depósito: 01/04/2010

(45) Data de Concessão: 24/11/2020

(54) Título: APARELHO E PROCESSO PARA A GERAÇÃO DE UM SINAL DE ÁUDIO DE SÍNTESE E PARA A CODIFICAÇÃO DE UM SINAL DE ÁUDIO

(51) Int.Cl.: G10L 19/008; G10L 19/18; G10L 21/038.

(52) CPC: G10L 19/008; G10L 19/18; G10L 21/038.

(30) Prioridade Unionista: 09/04/2009 US 61/168,068; 30/12/2009 EP 09181008.5.

(73) Titular(es): FRAUNHOFER-GESELLSCHAFT ZUR FÖRDERUNG DER ANGEWANDTEN FORSCHUNG E.V..

(72) Inventor(es): FREDERIK NAGEL; MARKUS MULTRUS; JÉRÉMIE LECOMTE; STEFAN.BAYER; GUILLAUME FUNCHS; JOHANNES HILPERT; JULIEN ROBILLIARD.

(86) Pedido PCT: PCT EP2010054434 de 01/04/2010

(87) Publicação PCT: WO 2010/115845 de 14/10/2010

(85) Data do Início da Fase Nacional: 12/05/2011

(57) Resumo: APARELHO E PROCESSO PARA A GERAÇÃO DE UM SINAL DE ÁUDIO DE SÍNTESE E PARA A CODIFICAÇÃO DE UM SINAL DE ÁUDIO Um aparelho para geração de sinal de áudio de síntese usando um sinal de controlo de patching compreende um primeiro conversor, um gerador de patch no domínio espectral, um manipulador de reconstrução de alta frequência e um combinador. O primeiro conversor é configurado para converter uma porção do tempo de um sinal de áudio em uma representação espectral.

"APARELHO E PROCESSO PARA A GERAÇÃO DE UM SINAL
DE ÁUDIO DE SÍNTESE E PARA A CODIFICAÇÃO DE UM SINAL DE ÁUDIO"

Descrição

A presente invenção se refere a um processamento
5 de sinal de áudio, e em particular, a um aparelho e um método para
a geração de um sinal de áudio de síntese, um aparelho e um método
para a codificação de um sinal de áudio e de um sinal de áudio
codificado.

A armazenagem e a transmissão de sinais de áudio
10 estão geralmente sujeitas a rígidas restrições de taxa de bits.
Essas restrições são geralmente superadas por uma codificação
intermediária do sinal. No passado, os codificadores eram forçados
a reduzir drasticamente a largura de banda do áudio transmitido
quando somente era disponível uma taxa muito baixa de bits. Os
15 modernos codécs de áudio podem codificar sinais de banda larga
usando métodos de extensão de largura de banda (BWE), como
descrito em *M. Dietz, L. Liljeryd, K. Kjörling and O. Kunz,*
"Spectral Band Replication, a novel approach in audio coding" in
112th AES Convention, Munich, May 2002; S. Meltzer, R. Böhm and F.
20 *Henn, "SBR enhanced audio codecs for digital broadcasting such as*
"Digital Radio Mondiale" (DRM)," in 112th AES Convention, Munich,
May 2002; T. Ziegler, A. Ehret, P. Ekstrand and M. Lutzky,
"Enhancing mp3 with SBR: Features and Capabilities of the new
mp3PRO Algorithm," in 112th AES Convention, Munich, May 2002;
25 *International Standard ISO/IEC 14496-3:2001/FPDAM 1, "Bandwidth*
Extension," ISO/IEC, 2002. Speech bandwidth extension method and
apparatus Vasu Iyengar et al. US Patent 5,455,888; E. Larsen, R.
M. Aarts, and M. Danessis. Efficient high-frequency bandwidth

- extension of music and speech. In AES 112th Convention, Munich, Germany, May 2002; R.M. Aarts, E. Larsen, and O. Ouweltjes. A unified approach to low-and high frequency bandwidth extension. In AES 115th Convention, New York, USA, October 2003; K. Käyhkö. A
- 5 Robust Wideband Enhancement for Narrowband Speech Signal. Research Report, Helsinki University of Technology, Laboratory of Acoustics and Audio Signal Processing, 2001; E. Larsen and R.M. Aarts. Audio Bandwidth Extension - Application to psychoacoustics, Signal Processing and Loudspeaker Design. John Wiley & Sons, Ltd, 2004;
- 10 E. Larsen, R.M. Aarts, and M. Danessis. Efficient high-frequency bandwidth extension of music and speech. In AES 112th Convention, Munich, Germany, May 2002; J. Makhoul. Spectral Analysis of Speech by Linear Prediction. IEEE Transactions of Audio and Electroacoustics, AU-21(3), June 1973; United States Patent
- 15 Application 08/951,029, Ohmori, et al. Audio band width extending system and method; United States Patent 6895375, Malah, D & Cox, R.V.: System for bandwidth extension of Narrow-band speech, and Frederik Nagel, Sascha Disch, "A harmonic bandwidth extension method for audio codecs," ICASSP International Conference on
- 20 Acoustics, Speech and Signal Processing, IEEE CNF, Taipei, Taiwan, April 2009.

Esses algoritmos se baseiam em uma representação paramétrica do conteúdo de alta frequência (HF). Esta representação é gerada a partir da parte de baixa frequência (LF)

25 do sinal decodificado por meio da transposição na região espectral HF ("patching") e a aplicação de um parâmetro acionado pós-processamento.

Na técnica, são usados os métodos de extensão de

largura de banda como a replicação de banda espectral (SBR) como um método eficiente para a geração de sinais de alta frequência em um codec de base HFR (reconstrução de alta frequência).

A replicação de banda espectral (SBR), como descrita em *M. Dietz, L. Liljeryd, K. Kjörling and O. Kunz, "Spectral Band Replication, a novel approach in audio coding" in 112th AES Convention, Munich, May 2002*, usa um banco de filtro em espelho de quadratura (QMF) para a geração das informações HF. Com o denominado "patching", são copiados sinais de banda QMF mais baixos em bandas QMF mais altas, levando à uma replicação das informações da parte LF na parte HF. A parte HF gerada é posteriormente adaptada à parte HF original com a ajuda de parâmetros que ajustam o envelope espectral e a tonalidade.

Em SBR, como padronizada em HE-AAC, todas as operações que incluem o patching por meio de cópia simples, são sempre feitas dentro do domínio QMF. Entretanto, podem ser feitos outros métodos diferentes de patching em diferentes domínios, como o domínio FFT ou o domínio do tempo. É possível imaginar a habilitação de SBR para escolher de forma alternativa um algoritmo de patching que opere tanto no domínio FFT como no domínio do tempo, e que precise de uma transformação adicional para a alimentação da etapa de análise QMF.

Em puro SBR, somente é disponível um algoritmo de patching que não leve em conta as necessidades de hardware ou de software especiais nem características de sinais. Assim, o SBR não pode adaptar o algoritmo de patching. Pode-se imaginar a simples escolha entre dois distintos algoritmos de patching. Como os dois métodos de patching funcionam em domínios diferentes, as áreas de

transição podem produzir problemas de bloqueio, o que torna a comutação de grãos finos entre os dois métodos praticamente impossível.

5 A WO 98/57436 revela os métodos de transposição usados na replicação de banda espectral, que são combinados com o ajuste do envelope espectral.

A WO 02/052545 ensina que os sinais podem ser classificados tanto de forma tipo trem de pulsos como a não tipo trem de pulsos e, baseado nessa classificação, é proposto um transposer de comutação adaptativa. O transposer de comutação realiza dois algoritmos de patching em paralelo e a unidade de mixagem combina ambos os sinais patched dependentes da classificação (trem de pulsos ou não trem de pulsos). A comutação real entre ou a mixagem dos transposers é feita em um banco de
10 filtros de ajuste de envelope em resposta aos dados de envelope e controle. Além disso, para sinais do tipo trem de pulsos, o sinal base é transformado em um domínio de banco de filtros, é feita uma operação de translação de frequências realizada em um ajuste de envelope do resultado da translação de frequências. Este é um
15 procedimento de patching combinado/processamento adicional. Para sinais do tipo não trem de pulsos, é provido um transposer no domínio da frequência (transposer FD) e o resultado do transposer no domínio da frequência é então transformado no domínio de banco de filtros, onde é feito o ajuste de envelope. Assim, a
20 implementação e a flexibilidade deste procedimento, que tem em uma alternativa, uma abordagem de patching combinado/processamento adicional, e que tem na outra alternativa, o transposer no domínio da frequência, que é posicionado fora do banco de filtros onde
25

ocorre o ajuste de envelope, é problemático com relação às possibilidade de flexibilidade e implementação.

É um objetivo da presente invenção prover um conceito para a geração de um sinal de áudio de síntese permitindo
5 uma melhor qualidade e uma eficiente implementação.

Este objetivo é alcançado por meio de um aparelho para a geração de um sinal de áudio de síntese de acordo com a reivindicação 1, um aparelho para a codificação de um sinal de áudio de acordo com a reivindicação 10, um método para a geração
10 de acordo com a reivindicação 12, um método para a codificação de acordo com a reivindicação 13, um sinal de áudio codificado de acordo com a reivindicação 14 ou um programa de computador de acordo com a reivindicação 15.

A presente invenção se sustenta na idéia básica
15 de que a recém mencionada melhor qualidade e/ou a implementação eficiente podem ser obtidas quando uma porção do tempo de um sinal de áudio é convertida em uma representação espectral antes de realizar uma pluralidade de diferentes algoritmos de patching no domínio espectral, em que cada algoritmo de patching gera uma
20 representação espectral modificada compreendendo componentes espectrais em uma banda superior de frequência obtida a partir dos componentes espectrais correspondentes em uma banda de frequência núcleo do sinal de áudio, e selecionando um primeiro algoritmo de patching no domínio espectral a partir de uma pluralidade de
25 algoritmos de patching de uma primeira porção do tempo e um segundo algoritmo de patching no domínio espectral da pluralidade de algoritmos de patching de uma segunda porção do tempo diferente de acordo com um sinal de controle de patching para obter a

representação espectral modificada. Com essa medida, pode ser evitada a qualidade e/ou flexibilidade reduzidas devido à comutação entre dois algoritmos de patching em domínios diferentes e, portanto o processamento pode ser menos complexo, mantendo a
5 qualidade de percepção.

De acordo com uma configuração da presente invenção, um aparelho para a geração de um sinal de áudio de síntese usando um sinal de controle de patching compreende um primeiro conversor, um gerador de patch no domínio espectral, um
10 manipulador de reconstrução de alta frequência e um combinador. O primeiro conversor é configurado para converter uma porção do tempo de um sinal de áudio em uma representação espectral. O gerador de patch no domínio espectral é configurado para realizar uma pluralidade de diferentes algoritmos de patching no domínio
15 espectral, onde cada algoritmo de patching gera uma representação espectral modificada compreendendo componentes espectrais em uma banda superior de frequência obtida a partir dos componentes espectrais correspondentes em uma banda de frequência núcleo do sinal de áudio. O gerador de patch no domínio espectral é também
20 configurado para selecionar um primeiro algoritmo de patching no domínio espectral a partir da pluralidade de algoritmos de patching de uma primeira porção do tempo e um segundo algoritmo de patching no domínio espectral a partir da pluralidade de algoritmos de patching da segunda porção do tempo diferente de
25 acordo com o sinal de controle de patching para a obtenção da representação espectral modificada. O manipulador de reconstrução de alta frequência é configurado para manipular a representação espectral modificada ou um sinal obtido a partir da representação

espectral modificada de acordo com o parâmetro de replicação de banda espectral para obter um sinal ampliado de largura de banda. O combinador é configurado para combinar o sinal de áudio tendo componentes espectrais na banda de frequência núcleo ou um sinal obtido a partir do sinal de áudio com o sinal ampliado de largura de banda para obter o sinal de áudio de síntese.

De acordo com outra configuração da presente invenção, um aparelho para a codificação de um sinal de áudio compreende um codificador núcleo, um extrator paramétrico e um calculador paramétrico. O sinal de áudio compreende uma banda de frequência núcleo e uma banda superior de frequência. O codificador núcleo é configurado para a codificação do sinal de áudio dentro da banda de frequência núcleo. O extrator paramétrico é configurado para a extração de um sinal de controle de patching do sinal de áudio, o sinal de controle de patching indicando um algoritmo de patching selecionado a partir da pluralidade de diferentes algoritmos de patching no domínio espectral, o algoritmo de patching selecionado a ser feito em um domínio espectral para a geração de um sinal de áudio de síntese em um decodificador de extensão de largura de banda. O calculador paramétrico é configurado para calcular o parâmetro de replicação de banda espectral da banda superior de frequência.

De acordo com outra configuração, um fluxo de dados de sinal de áudio codificado compreende um sinal de áudio codificado no interior de uma banda de frequência núcleo, um sinal de controle de patching, o sinal de controle de patching indicando um algoritmo de patching selecionado a partir de uma pluralidade de diferentes algoritmos de patching no domínio espectral, o

algoritmo de patching selecionado a ser feito no domínio espectral para a geração de um sinal de áudio de síntese em um decodificador de extensão de largura de banda e um parâmetro de replicação de banda espectral calculado a partir de uma banda superior de
5 frequência do sinal de áudio.

Portanto, as configurações da presente invenção se referem a um conceito para a comutação entre pelo menos dois diferentes algoritmos de patching no domínio espectral de um grupo de algoritmos de patching no domínio espectral. O grupo de
10 algoritmos de patching pode compreender um primeiro algoritmo de patching compreendendo uma transposição harmônica baseada em um vocoder de fase simples e de funcionalidades SBR de cópia não harmônica, um segundo algoritmo de patching compreendendo uma transposição harmônica baseada em um vocoder de fase múltipla, um
15 terceiro algoritmo de patching compreendendo funcionalidades SBR de cópia não harmônica e um quarto algoritmo de patching compreendendo uma distorção não linear. Além disso, a extensão de largura de banda pode ser feita de maneira que o sinal ampliado de largura de banda compreenda a banda superior de frequência tendo
20 uma frequência máxima de pelo menos quatro vezes a frequência de cruzamento na banda de frequência núcleo.

Como resultado, pela comutação entre os pelo menos dois diferentes algoritmos de patching no domínio espectral, pode ser obtida uma complexidade reduzida com a mesma qualidade de
25 percepção como dentro de um cenário de extensão de largura de banda.

Outras configurações da presente invenção se referem a um aparelho não compreendendo um transformador

tempo/frequência para a transformação de um sinal de domínio do tempo obtido a partir de uma representação espectral modificada no domínio espectral. Portanto, as configurações permitem que o manipulador de reconstrução de alta frequência possa operar em uma
5 representação espectral modificada diretamente sem exigir outra transformação (por exemplo, uma análise QMF) do domínio do tempo para o domínio espectral, como no caso de uma abordagem de patching combinado/processamento adicional sendo operativa nos domínios diferentes.

10 Outras configurações da presente invenção se referem a um extrator paramétrico que é configurado para determinar a partir da pluralidade de diferentes algoritmos de patching no domínio espectral um algoritmo de patching selecionado. Aqui, o algoritmo de patching selecionado se baseia
15 em uma comparação do sinal de áudio ou de um sinal obtido a partir do sinal de áudio com uma pluralidade de sinais ampliados de largura de banda tendo sido obtidos pela realização da pluralidade de algoritmos de patching no domínio espectral e a manipulação de uma representação espectral modificada de uma porção do tempo do
20 sinal de áudio. Portanto, as configurações provêem um método de seleção do algoritmo de patching ideal para a geração de um sinal de áudio de síntese em um decodificador de extensão de largura de banda.

Podem ser usados parâmetros de controle para
25 decidir qual patching é o mais adequado. Para obter isso, pode ser usado um estágio de análise por síntese; isto é, todos os patches podem ser aplicados, sendo escolhido o melhor de acordo com um objetivo. No modo preferido da invenção, o objetivo é obter a

melhor qualidade de percepção da restituição. Em modos alternativos, uma função do objetivo deve ser otimizada. Por exemplo, o objetivo pode ser o de preservar a planicidade espectral dos HFs originais da forma mais verdadeira possível.

5 Por um lado, a seleção de patching pode ser feita somente no codificador considerando o sinal original, o sinal sintetizado ou ambos. A decisão (sinal de controle de patching) é então transmitida ao decodificador. Por outro lado, a seleção pode ser feita de forma síncrona nos lados do codificador e do
10 decodificador considerando somente a largura de banda núcleo do sinal sintetizado. O último método não precisa gerar uma informação auxiliar adicional.

A seguir, são esclarecidas as configurações da presente invenção com referência aos desenhos de acompanhamento,
15 onde:

A Fig. 1a mostra um diagrama de blocos de uma configuração de um aparelho para a geração de um sinal de áudio de síntese usando um sinal de controle de patching;

A Fig. 1b mostra um diagrama de blocos de uma
20 implementação de um gerador de patch no domínio espectral da Fig. 1a;

A Fig. 2a mostra um diagrama de blocos de outra configuração de um aparelho para a geração de um sinal de áudio de síntese;

25 A Fig. 2b mostra uma ilustração esquemática de um esquema de extensão de largura de banda;

A Fig. 3 mostra uma ilustração esquemática de um primeiro algoritmo de patching exemplar;

A Fig. 4 mostra uma ilustração esquemática de um segundo algoritmo de patching exemplar;

A Fig. 5 mostra uma ilustração esquemática de um terceiro algoritmo de patching exemplar;

5 A Fig. 6 mostra uma ilustração esquemática de um quarto algoritmo de patching exemplar;

A Fig. 7 mostra um diagrama de blocos de uma configuração da Fig. 1a sem um transformador tempo/frequência colocado depois do gerador de patch no domínio espectral;

10 A Fig. 8 mostra um diagrama de blocos de uma configuração da Fig. 1a com um segundo conversor (conversor frequência/tempo);

A Fig. 9 mostra um diagrama de blocos de uma configuração de um aparelho para a codificação um sinal de áudio;

15 A Fig. 10 mostra um diagrama de blocos de outra configuração de um aparelho para a codificação de um sinal de áudio; e

20 A Fig. 11 mostra uma visão geral de uma configuração de um esquema de patching em um domínio da frequência.

A Fig. 1a mostra um diagrama de blocos de um aparelho 100 para a geração de um sinal de áudio de síntese 145 usando um sinal de controle de patching 119 de acordo com uma configuração da presente invenção. O aparelho 100 compreende um primeiro conversor 110, a gerador de patch no domínio espectral 25 120, um manipulador de reconstrução de alta frequência 130 e um combinador 140. O primeiro conversor 110 é configurado para converter uma porção do tempo de um sinal de áudio 105 em uma

representação espectral 115. O gerador de patch no domínio espectral 120 é configurado para realizar uma pluralidade 117-1 de diferentes algoritmos de patching no domínio espectral, em que cada algoritmo de patching gera uma representação espectral modificada 125 compreendendo componentes espectrais em uma banda superior de frequência obtida a partir dos componentes espectrais correspondentes em uma banda de frequência núcleo do sinal de áudio 105. Como mostrado na Fig. 1b, o gerador de patch no domínio espectral 120 pode ser configurado para selecionar um primeiro algoritmo de patching no domínio espectral 117-2 a partir da pluralidade 117-1 de algoritmos de patching de uma primeira porção do tempo 107-1 e um segundo algoritmo de patching no domínio espectral 117-3 a partir da pluralidade 117-1 de algoritmos de patching de uma segunda porção do tempo diferente 107-2 de acordo com o sinal de controle de patching 119 para obter uma representação espectral modificada 125.

O manipulador de reconstrução de alta frequência 130 é configurado para manipular a representação espectral modificada 125 ou um sinal obtido a partir da representação espectral modificada 125 de acordo com um parâmetro de replicação de banda espectral 127 para a obtenção de um sinal ampliado de largura de banda 135. O sinal obtido a partir da representação espectral modificada 125 pode, por exemplo, ser um sinal em um domínio QMF tendo sido obtido após a aplicação de uma análise QMF a um sinal de domínio do tempo modificado baseando-se em uma representação espectral modificada 125. O combinador 140 é configurado para combinar o sinal de áudio 105 tendo componentes espectrais na banda de frequência núcleo ou um sinal obtido a

partir do sinal de áudio 105 com o sinal ampliado de largura de banda 135 para obter o sinal de áudio de síntese 145. Aqui, o sinal obtido a partir do sinal de áudio 105 pode, por exemplo, ser um sinal de baixa frequência decodificado que tenha sido obtido
5 após a decodificação de um sinal de áudio codificado dentro da banda de frequência núcleo.

Como pode ser visto na Fig. 1a, o gerador de patch no domínio espectral 120 do aparelho 100 é implementado para operar em um domínio espectral e não em um domínio do tempo.

10 A Fig. 2a mostra um diagrama de blocos de outra configuração de um aparelho 200 para a geração do sinal de áudio de síntese 145. Aqui, os componentes do aparelho 200 da Fig. 2a, que são os mesmos do aparelho 100 da Fig. 1a, são omitidos e não mostrados ou descritos novamente. Na configuração como mostrada na
15 Fig. 2a, o gerador de patch no domínio espectral 120 do aparelho 200 é configurado para realizar pelo menos dois diferentes algoritmos de patching no domínio espectral a partir de um grupo 203 de algoritmos de patching no domínio espectral. O grupo 203 de algoritmos de patching compreende um primeiro algoritmo de
20 patching 205-1 compreendendo uma transposição harmônica com base em um vocóder de fase simples e funcionalidades SBR de cópia não harmônica, um segundo algoritmo de patching 205-2 compreendendo uma transposição harmônica com base em um vocóder de fase múltipla, um terceiro algoritmo de patching 205-3 compreendendo
25 funcionalidades SBR de cópia não harmônica e um quarto algoritmo de patching 205-4 compreendendo uma distorção não linear.

Como mostrado na Fig. 2b, o aparelho 200 pode ser adaptado para fazer uma extensão de largura de banda de maneira

que o sinal ampliado de largura de banda 135 compreenda a banda superior de frequência 220 tendo uma frequência máxima 225 de pelo menos quatro vezes a frequência de cruzamento 215 na banda de frequência núcleo 210. No contexto de SBR, o valor típico da

5 frequência de cruzamento 215 definida como a maior frequência da banda de frequência núcleo 210 pode, por exemplo, estar em uma faixa abaixo de 4 kHz, 5 kHz ou 6 kHz. Conseqüentemente, a frequência máxima 225 da banda superior de frequência 220 pode, por exemplo, ser cerca de 16 kHz, 20 kHz ou 24 kHz.

10 A Fig. 3 mostra uma ilustração esquemática de um primeiro algoritmo de patching exemplar 205-1. Em particular, o gerador de patch no domínio espectral 120 é configurado para realizar um algoritmo de patching selecionado a partir de pelo menos dois diferentes algoritmos de patching no domínio espectral,

15 o algoritmo de patching selecionado compreendendo o primeiro algoritmo de patching 205-1. O primeiro algoritmo de patching 205-1 compreende a transposição harmônica com base em um vocoder de fase simples 305 compreendendo um fator de extensão de largura de banda(σ) de dois, controlando uma transformada de uma banda de

20 frequência fonte 310 extraída da banda de frequência núcleo 210 em uma primeira banda de frequência alvo 310'. Aqui, as fases dos componentes espectrais na banda de frequência fonte 310 são multiplicadas pelo fator de extensão de largura de banda(σ) de maneira que a primeira banda de frequência alvo 310 tenha

25 frequências que variem da frequência de cruzamento (f_x) até o dobro da frequência de cruzamento (f_x). O primeiro algoritmo de patching 205-1 ainda compreende funcionalidades SBR de cópia não harmônica 315 para a transformação dos componentes espectrais na

primeira banda de frequência alvo 310' em uma segunda banda de frequência alvo 320' por uma primeira cópia, de maneira que a segunda banda de frequência alvo 320' tenha frequências que variem do dobro da frequência de cruzamento (f_x) a três vezes a frequência de cruzamento (f_x) e para outras transformações dos componentes espectrais na segunda banda de frequência alvo 320' em uma terceira banda de frequência alvo 330' por uma segunda cópia de maneira que a terceira banda de frequência alvo 330' tenha frequências que variem de três vezes a frequência de cruzamento (f_x) a quatro vezes a frequência de cruzamento (f_x) incluída na banda superior de frequência 220, a banda superior de frequência 220 compreendendo a primeira 310', a segunda 320' e a terceira 330' bandas de frequência alvo. Em particular, como mostrado na Fig. 3, o sinal ampliado de largura de banda 135 compreende a banda superior de frequência 220 gerada a partir da banda de frequência núcleo 210, onde a banda superior de frequência 220 tem uma frequência máxima de quatro vezes a frequência de cruzamento (f_x).

A Fig. 4 mostra uma ilustração esquemática de um segundo algoritmo de patching exemplar 205-2. Aqui, em particular, o gerador de patch no domínio espectral 120 é configurado para realizar um algoritmo de patching selecionado a partir de pelo menos dois diferentes algoritmos de patching no domínio espectral, o algoritmo de patching selecionado compreendendo o segundo algoritmo de patching 205-2. O segundo algoritmo de patching 205-2 compreende uma transposição harmônica com base em um vocoder de fase múltipla 405 compreendendo um primeiro fator de extensão de largura de banda (σ_1) de 2 controlando uma transformada de uma

primeira banda de frequência fonte 410 extraída da banda de frequência núcleo 210 em uma primeira banda de frequência alvo 410'. Aqui, as fases dos componentes espectrais na primeira banda de frequência fonte 410 são multiplicadas pelo primeiro fator de extensão de largura de banda (σ_1), de maneira que a primeira banda de frequência alvo 410' tenha frequências que variem da frequência de cruzamento (f_x) até o dobro da frequência de cruzamento (f_x). O segundo algoritmo de patching 205-2 ainda compreende um segundo fator de extensão de largura de banda (σ_2) de 3, controlando uma transformada de uma segunda banda de frequência fonte 420-1, 420-2 extraída da banda de frequência núcleo 210 em uma segunda banda de frequência alvo 420', 420''. Aqui, as fases dos componentes espectrais na segunda banda de frequência fonte 420-1, 420-2 são multiplicadas pelo segundo fator de extensão de largura de banda (σ_2) de maneira que a segunda banda de frequência alvo 420', 420'' tenha frequências que variem do dobro da frequência de cruzamento (f_x) a três vezes a frequência de cruzamento (f_x) ou variando entre a frequência de cruzamento (f_x) a três vezes a frequência de cruzamento (f_x), respectivamente. Finalmente, o segundo algoritmo de patching 205-2 ainda compreende um terceiro fator de extensão de largura de banda (σ_3) de 4, controlando uma transformada de uma terceira banda de frequência fonte 430-1, 430-2 extraída da banda de frequência núcleo 210 em uma terceira banda de frequência alvo 430', 430''. Aqui, as fases dos componentes espectrais na terceira banda de frequência fonte 430-1, 430-2 são multiplicadas pelo terceiro fator de extensão de largura de banda (σ_3) de maneira que a terceira banda de frequência alvo 430', 430'' tenha frequências que variem de três vezes a frequência de

cruzamento (f_x) a quatro vezes a frequência de cruzamento (f_x) ou variando entre a frequência de cruzamento (f_x) a quatro vezes a frequência de cruzamento (f_x) incluída na banda superior de frequência 220, respectivamente. Como no primeiro algoritmo de patching 205-1 mostrado na Fig. 3, a banda superior de frequência 220 do sinal ampliado de largura de banda 135 compreende a primeira 410', a segunda 420', 420'' e a terceira 430', 430'' bandas de frequência alvo tendo uma frequência máxima de quatro vezes a frequência de cruzamento (f_x).

10 A Fig. 5 mostra uma ilustração esquemática de um terceiro algoritmo de patching exemplar 205-3. Na configuração da Fig. 5, o gerador de patch no domínio espectral 120 é configurado para realizar um algoritmo de patching selecionado a partir de pelo menos dois diferentes algoritmos de patching no domínio

15 espectral, o algoritmo de patching selecionado compreendendo o terceiro algoritmo de patching 205-3. O terceiro algoritmo de patching 205-3 compreende funcionalidades SBR de cópia não harmônica 505 para a transformação dos componentes espectrais em uma banda de frequência fonte 510 sendo a banda de frequência

20 núcleo 210 em uma banda de frequência alvo 510' por uma primeira cópia de maneira que a primeira banda de frequência alvo 510' tenha frequências que variem da frequência de cruzamento (f_x) até o dobro da frequência de cruzamento (f_x). Componentes espectrais na primeira banda de frequência alvo 510' ainda são transformados

25 em uma segunda banda de frequência alvo 520' por uma segunda cópia, de maneira que a segunda banda de frequência alvo 520' tenha frequências que variem do dobro da frequência de cruzamento (f_x) a três vezes a frequência de cruzamento (f_x). Finalmente,

componentes espectrais na segunda banda de frequência alvo 520' são ainda transformados em uma terceira banda de frequência alvo 530' por uma terceira cópia, de maneira que a terceira banda de frequência alvo 530' tenha frequências que variem de três vezes a frequência de cruzamento (f_x) a quatro vezes a frequência de cruzamento (f_x) incluída na banda superior de frequência 220. Outra vez, a banda superior de frequência 220 do sinal ampliado de largura de banda 135 compreende a primeira 510', segunda 520' e terceira 530' bandas de frequência alvo tendo uma frequência máxima de quatro vezes a frequência de cruzamento (f_x).

A Fig. 6 mostra uma ilustração esquemática de um quarto algoritmo de patching exemplar 205-4. Na configuração da Fig. 6, o gerador de patch no domínio espectral 120 é configurado para realizar um algoritmo de patching selecionado a partir de pelo menos dois diferentes algoritmos de patching no domínio espectral, o algoritmo de patching selecionado compreendendo o quarto algoritmo de patching 205-4. Aqui, o quarto algoritmo de patching 205-4 compreende uma distorção não linear para a geração de componentes espectrais na banda superior de frequência 220 tendo frequências variando da frequência de cruzamento (f_x) a quatro vezes a frequência de cruzamento (f_x).

Em geral, nas configurações das Figs. 3-6 como descritas acima, os algoritmos de patching no domínio espectral 205-1; 205-2; 205-3; 205-4 são realizados com o gerador de patch no domínio espectral 120 sendo configurado para transformar uma componente espectral em uma banda inicial 310, 310', 320'; 410, 420-1, 420-2, 430-1, 430-2; 510, 510', 520' obtida a partir da banda de frequência núcleo 210 ou de uma banda superior de

frequência não incluída na banda de frequência núcleo 210 em uma componente espectral alvo na banda superior de frequência 220 de maneira que a componente espectral alvo seja diferente para cada um dos algoritmos de patching no domínio espectral.

5 Em particular, o gerador de patch no domínio espectral 120 pode compreender um filtro de passagem de banda para a extração da banda inicial da banda de frequência núcleo 210 ou uma banda superior de frequência 220, em que uma característica de
10 passagem de banda do filtro de passagem de banda possa ser selecionada, de maneira que a banda inicial seja transformada nas bandas de frequência alvo correspondentes 310', 320', 330'; 410', 420', 420'', 430', 430''; 510', 520', 530' como mostrado na Figs. 3-6.

 Os diferentes algoritmos de patching no domínio
15 espectral 205-1; 205-2; 205-3; 205-4 podem ser realizados de acordo com um desempenho exigido, como dentro do esquema de extensão de largura de banda da Fig. 2b.

 Especificamente, empregando um vocoder de fase
20 simples ou de fase múltipla como mostrado por exemplo na Fig. 3 ou na Fig. 4, respectivamente, a estrutura da frequência é harmônica e corretamente ampliada no domínio da alta frequência, porque a banda base (por exemplo, a banda de frequência núcleo 210) é espectralmente dispersa por um múltiplo par (por exemplo, $\sigma_1 = 2$, $\sigma_2 = 3$, $\sigma_3 = 4$), e porque os componentes espectrais na banda base
25 são combinados com os componentes espectrais adicionais gerados.

 Um algoritmo de patching baseado em vocoder de fase pode ser vantajoso se a banda base já estiver fortemente limitada na largura de banda, por exemplo, usando somente uma taxa

de bits muito baixa. Assim, a reconstrução dos componentes da frequência superior já começa em uma frequência relativamente baixa. A frequência de cruzamento típica é, nesse caso, menor que cerca de 5 KHz (ou até menor que 4 KHz). Nessa região, o ouvido humano é muito sensível a dissonâncias, devido aos harmônicos incorretamente posicionados. Isto pode resultar em uma impressão de tons "não naturais". Além disso, tons com espaçamentos espectralmente próximos (com a dissonância espectral de cerca de 30 Hz a 300 Hz) são percebidos como tons grosseiros. A continuação harmônica da estrutura da frequência da banda base evita essas impressões auditivas incorretas e desagradáveis.

Além disso, empregando as funcionalidades SBR de cópia não harmônica como mostrado, por exemplo na Fig. 5, as regiões espectrais podem ser copiadas como subbandas a uma região de maior frequência ou na região de frequência a ser replicada. Também a cópia se baseia na observação, o que é verdade para todos os métodos de patching, que as propriedades espectrais dos sinais de maior frequência sejam similares em muitos aspectos às propriedades dos sinais da banda base. Existem somente poucos desvios entre eles. Além disso, o ouvido humano não é tipicamente muito sensível à alta frequência (tipicamente iniciando em cerca de 5 KHz), especialmente em relação a um mapeamento espectral não preciso. Realmente, esta é em geral a idéia principal da replicação de banda espectral. Copiar em particular, compreende a vantagem de que é fácil e rápido para implementar. Este algoritmo de patching também tem uma alta flexibilidade em relação aos limites do patch, já que a cópia do espectro pode ser feita em qualquer limite de subbanda.

Finalmente, o algoritmo de patching de distorção não linear (ver, por exemplo, Fig. 6) pode compreender uma geração de harmônicos por corte, limitação, quadratura, etc. Se, por exemplo, um sinal difundido for ocupado espectralmente de forma muito fina (por exemplo, após a aplicação do algoritmo de patching do vocoder de fase acima mencionado), é possível que o espectro difundido possa ser opcional e adicionalmente suplementado por um sinal distorcido para evitar buracos indesejados de frequência.

Deve ser notado que além dos algoritmos de patching acima mencionados do grupo 203 de algoritmos de patching (ver Fig. 2a), podem ser realizados outros algoritmos de patching no domínio espectral como o espelhamento espectral.

Na configuração da Fig. 7, um aparelho 700 é mostrado não compreendendo um transformador tempo/frequência, como indicado pelo bloco tracejado 710 para a transformação de um sinal de domínio do tempo 705 obtida a partir da representação espectral modificada 125 no domínio espectral. Isto significa que, nesse caso, o manipulador de reconstrução de alta frequência 130 receberá em sua entrada uma representação espectral modificada 125 e não um sinal de domínio da frequência 715, presente na saída deste transformador tempo/frequência 710.

A configuração descrita pode ser vantajosa, porque nesse caso o novo processamento da representação espectral modificada 125 realizado pelo manipulador de reconstrução de alta frequência 130 pode ocorrer prontamente no mesmo domínio (por exemplo, no domínio FFT ou QMF) em que o algoritmo de patching realizado pelo gerador de patch no domínio espectral 120 é operativo. Portanto, não será necessária outra transformada entre

domínios diferentes como uma transformada do domínio do tempo para o domínio espectral (por exemplo, uma análise QMF), que leva a uma implementação mais fácil.

Na configuração da Fig. 8, é mostrado um aparelho 800 ainda compreendendo um segundo conversor 810 para a conversão de uma representação espectral modificada 125 no domínio do tempo. Novamente, são omitidos os componentes do aparelho 800 da Fig. 8, que podem corresponder aos do aparelho 100 da Fig. 1a. Como mostrado na Fig. 8, o segundo conversor 810 pode ser adaptado para aplicar uma síntese combinada a uma análise aplicada pelo primeiro conversor 110. Aqui, o primeiro conversor 110 é configurado para fazer uma conversão tendo um primeiro comprimento de conversão 111, enquanto o segundo conversor 810 é configurado para fazer uma conversão tendo um segundo comprimento de conversão. Em particular, o segundo comprimento de conversão pode depender de uma característica da extensão de largura de banda onde são levados em conta a razão da frequência máxima (F_{max}) na banda superior de frequência 220 e a frequência de cruzamento (f_c) na banda de frequência núcleo 210 e o primeiro comprimento de conversão 111.

Nas configurações da presente invenção, o primeiro conversor 110 pode, por exemplo, ser implementado para realizar uma transformada rápida de Fourier (FFT), uma transformada breve de Fourier (STFT), uma transformada discreta de Fourier (DFT) ou uma análise QMF, enquanto o segundo conversor 810 pode, por exemplo, ser implementado para realizar uma transformada inversa rápida de Fourier (IFFT), uma transformada inversa breve de Fourier (ISTFT), uma transformada inversa discreta de Fourier

(IDFT) ou uma síntese QMF.

Especificamente, o segundo comprimento de conversão pode ser escolhido de maneira que seja igual à razão f_{\max}/f_x multiplicada pelo primeiro comprimento de conversão 111. Assim, o segundo comprimento de conversão ou resolução de frequência aplicada ao segundo conversor 810 será prontamente adaptada à característica da extensão de largura de banda do esquema de extensão de largura de banda como mostrado na Fig. 2b. Isto ocorre uma vez que a característica da extensão de largura de banda é essencialmente regida pela razão acima (f_{\max}/f_x) correspondendo a uma maior taxa efetiva de amostragem de acordo com o princípio de Nyquist.

A Fig. 9 mostra um diagrama de blocos de uma configuração de um aparelho 900 para a codificação de um sinal de áudio 105. O sinal de áudio 105 compreende uma banda de frequência núcleo 210 e uma banda superior de frequência 220. Em particular, o aparelho 900 para a codificação compreende um codificador núcleo 910, um extrator paramétrico 920 e um calculador paramétrico 930. O codificador núcleo 910 é configurado para a codificação do sinal de áudio 105 dentro da banda de frequência núcleo 210 para obter um sinal de áudio codificado 915 codificado dentro da banda de frequência núcleo 210. Além disso, o extrator paramétrico 920 é configurado para extrair um sinal de controle de patching 119 do sinal de áudio 105, o sinal de controle de patching 119 indicando um algoritmo de patching selecionado a partir da pluralidade 117-1 de diferentes algoritmos de patching no domínio espectral. Especificamente, o algoritmo de patching selecionado pode ser realizado em um domínio espectral para a geração do sinal de áudio

de síntese em um decodificador de extensão de largura de banda. Finalmente, o calculador paramétrico 930 é configurado para calcular um parâmetro SBR 127 da banda superior de frequência 220. O parâmetro SBR 127 calculado da banda superior de frequência 220, o sinal de controle de patching 119 indicando o algoritmo de patching selecionado e o sinal de áudio codificado 915 codificado dentro da banda de frequência núcleo 210 podem constituir um sinal de áudio codificado 935 a ser armazenado ou transmitido dentro de um fluxo de bits.

Na configuração da Fig. 9, o extrator paramétrico 920 pode ser configurado para analisar o sinal de áudio 105 ou um sinal obtido a partir do sinal de áudio 105 para determinar o sinal de controle de patching 119 com base em uma característica de sinal do sinal analisado. Por exemplo, o sinal de controle de patching 119 pode indicar um primeiro algoritmo de patching de uma primeira porção do tempo 107-1 do sinal analisado sendo caracterizado como 'fala' e um segundo algoritmo de patching para uma segunda porção do tempo 107-2 do sinal analisado sendo caracterizado como 'música estacionária'.

Assim, no caso de um sinal de fala, pode ser usado um processamento baseado em um modelo de origem de fala ou um modelo de geração de informações, como dentro de um domínio LPC (codificação preditiva linear), enquanto no caso da música estacionária, pode ser usado um modelo de origem estacionário ou um modelo de escoadouro de informações. Enquanto que no primeiro caso, é descrito o som que gera o sistema de fala humana/geração sonora, no último caso, é descrito o som que recebe o sistema de audição humano.

Além disso, pode ser implementado um esquema de processamento dependente do sinal pela comutação entre a transposição harmônica da porção do tempo compreendendo um evento transiente e uma operação de cópia não harmônica para a porção do tempo não compreendendo um evento transiente.

O procedimento acima correspondente a um loop aberto se baseia em uma análise direta do sinal de áudio 105 ou de um sinal obtido a partir do sinal de áudio 105 com relação à sua característica de sinal. Alternativamente, o extrator paramétrico 920 também pode operar em um loop fechado correspondente a uma implementação "análise por síntese".

Na configuração da Fig. 10, é ilustrado um aparelho 1000 para a codificação de um sinal de áudio 105 dentro de uma implementação análise por síntese. Especificamente, o extrator paramétrico 920 do aparelho 1000 para a codificação pode ser configurado para determinar a partir da pluralidade 117-1 de diferentes algoritmos de patching no domínio espectral do algoritmo de patching selecionado. Aqui, o algoritmo de patching selecionado pode se basear em uma comparação do sinal de áudio 105 ou de um sinal obtido a partir do sinal de áudio 105 com a pluralidade 1005 de sinais ampliados de largura de banda tendo sido obtidos pela realização da pluralidade 117-1 de algoritmos de patching no domínio espectral e a manipulação da representação espectral modificada 125 de uma porção do tempo do sinal de áudio 105. A comparação pode, por exemplo, ser feita por meio de uma unidade de seleção do algoritmo de patching 1010 calculando os parâmetros (SFM_{1005}) da medida de planicidade espectral (SFM) a partir da pluralidade 1005 de sinais ampliados de largura de banda

e do sinal de áudio 105 (SFM_{ref}), comparando os parâmetros SFM calculados SFM_{1005} e SFM_{ref} e selecionando a partir da pluralidade 117-1 de algoritmos de patching um algoritmo de patching específico (ideal), para o qual seja mínimo o desvio nos parâmetros SFM comparados. Finalmente, o algoritmo de patching ideal selecionado pode ser indicado pelo sinal de controle de patching 119 presente na saída do extrator paramétrico 920.

A Fig. 11 mostra uma visão geral de uma configuração de um esquema de patching em um domínio da frequência. Em particular, é mostrado um aparelho 1100 para a geração de um sinal ampliado de largura de banda como dentro do esquema de extensão de largura de banda da Fig. 2b. Na configuração da Fig. 11, o sinal de áudio 105 é representado por dados PCM 1101 (modulação de código de pulso) tendo um comprimento de frame de 1024 amostras ('frame: 1024'). Os dados PCM 1101 podem, por exemplo, ser um sinal de baixa frequência decodificado compreendendo uma banda base obtida a partir do sinal de áudio codificado 935, o sinal de áudio codificado 935 tendo sido transmitido de um aparelho para a codificação, como o codificador 900. Depois, pode ser usado um redutor da taxa de amostragem 1110 para fazer a redução da taxa de amostragem dos dados PCM 1101 por um fator de 2, por exemplo, para obter um sinal com taxa de amostragem reduzida 1115. O sinal com taxa de amostragem reduzida 1115 é ainda enviado a um janelador de análise 1120 indicado por um bloco denominado "janela" que pode ser configurado para gerar pluralidade de blocos consecutivos janelados de superposição de amostras de áudio. Aqui, cada bloco a partir da pluralidade de blocos consecutivos pode, por exemplo, compreender 512 amostras de

áudio. Além disso, uma primeira distância de tempo entre dois blocos consecutivos de amostras de áudio pode, por exemplo, ser ajustada para corresponder a 64 amostras indicadas por "Inc = 64". A sobreposição dos blocos consecutivos de amostras de áudio pode

5 também ser controlada selecionando uma função de janela de análise (ideal) adequada a partir da pluralidade de diferentes funções de janelas de análise aplicadas pelo janelador de análise 1120. A porção do tempo 1125 do sinal de áudio 105, que pode corresponder a um bloco consecutivo a partir da pluralidade de blocos

10 consecutivos de amostras de áudio, é ainda fornecido para o primeiro conversor 110 que pode ser implementado como, por exemplo, um processador FFT 1130 tendo o primeiro comprimento de conversão 111 de $N = 512$. O processador FFT 1130 pode ser configurado para converter a porção do tempo 1125 em uma

15 representação espectral 115 que pode, por exemplo, ser implementada sob forma polar 1135-1. Em particular, esta representação espectral 1135-1 compreende informações de magnitude 1135-2 e informações de fase 1135-3 que são ainda processadas por um gerador de patch no domínio espectral 1141, que pode

20 corresponder ao gerador de patch no domínio espectral 120 da Fig. 2a. O gerador de patch no domínio espectral 1141 da Fig. 11 pode compreender um primeiro algoritmo de patching 1141-1 indicado por "vocoder de fase mais cópia" correspondendo ao primeiro algoritmo de patching 205-1, um segundo algoritmo de patching 1143-1

25 indicado por "vocoder de fase" correspondendo ao segundo algoritmo de patching 205-2, um terceiro algoritmo de patching indicado por "função tipo SBR" correspondendo ao terceiro algoritmo de patching 205-3, e um quarto algoritmo de patching 1147-1 indicado por

"outra função, por exemplo, distorção não linear" correspondendo ao quarto algoritmo de patching 205-4 do grupo 203 de algoritmos de patching como mostrado na Fig. 2a.

Como descrito de forma correspondente no contexto da Fig. 2a anteriormente, o primeiro algoritmo de patching 1141-1 compreende um vocoder de fase simples 1141-2 e funcionalidades de cópia não harmônica 1141-3, 1141-4. Além disso, o segundo algoritmo de patching 1143-1, com base em uma operação de vocoder de fase múltipla compreende um primeiro vocoder de fase 1143-2, um segundo vocoder de fase 1143-3 e um terceiro vocoder de fase 1143-4. Além disso, o terceiro algoritmo de patching 1145-1 compreende funcionalidades SBR de cópia não harmônica realizando uma primeira operação de cópia 1145-2, uma segunda operação de cópia 1145-3 e uma terceira operação de cópia 1145-4. Finalmente, o quarto algoritmo de patching 1147-1 compreende uma funcionalidade de distorção não linear.

Particularmente, na configuração da Fig. 11, os sub-componentes dos blocos do algoritmo de patching 1141-1, 1143-1, 1145-1, 1147-1 podem corresponder àqueles blocos 205-1, 205-2, 205-3, 205-4 da Fig. 2a. Além disso, o símbolo ζ ('banda xover') pode corresponder à frequência de cruzamento (f_x).

Além disso, pode ser usado um seletor de patch 1150 para prover um sinal de controle de patching 1155 correspondendo ao sinal de controle de patching 119 para controlar o gerador de patch no domínio espectral 1141 de maneira que sejam realizados pelo menos dois diferentes algoritmos de patching no domínio espectral do grupo 1141-1, 1143-1, 1145-1, 1147-1 de algoritmos de patching, levando a uma representação espectral

modificada 1149 correspondendo a uma representação espectral modificada 125.

A representação espectral modificada 1149 pode ser (opcionalmente) processada por um interpolador subsequente 5 1160 para obter uma representação espectral interpolada modificada 1165. A representação espectral interpolada modificada 1165 pode então ser fornecida ao segundo conversor 810, que pode, por exemplo, ser implementado como um processador iFFT 1170 tendo um segundo comprimento de conversão de $N = 2048$. Aqui, como descrito 10 de forma correspondente na Fig. 8, o segundo comprimento de conversão de $N = 2048$ é ajustado para ser exatamente quatro vezes maior que o primeiro comprimento de conversão de $N = 512$. Assim, pode ser considerada a característica da extensão de largura de banda do esquema de extensão de largura de banda como realizada 15 com diferentes algoritmos de patching no domínio espectral, que foram explicados no detalhe anterior.

O processador iFFT 1170 pode ser configurado para converter a representação espectral interpolada modificada 1165 em um sinal modificado no domínio do tempo 1175 correspondendo ao 20 sinal modificado no domínio do tempo 815 da Fig. 8. O sinal modificado no domínio do tempo 1175 pode então ser enviado a um janelador de síntese 1180 para a aplicação de uma função de janela de síntese ao sinal modificado no domínio do tempo 1175 para obter um sinal modificado no domínio do tempo janelado 1185. Aqui, a 25 função de janela de síntese é combinada com a função de janela de análise de maneira que o efeito da aplicação da função de janela de análise é compensada pela aplicação da função de janela de síntese.

Como o sinal modificado no domínio do tempo janelado 1185 deve ser amostrado na maior taxa efetiva de amostragem (por exemplo, 32 KHz) conforme comparado com a razão de amostragem original (por exemplo, 8 KHz) devido à extensão de largura de banda, o sinal modificado no domínio do tempo janelado 1185 pode finalmente ser adicionado por sobreposição em um bloco 1190 indicado por "sobrepor e somar" já que a razão de uma segunda distância de tempo de, por exemplo, 256 amostras, indicada por "Inc = 256" aplicada ao bloco 1190 e a primeira distância de tempo de, por exemplo 64 amostras, aplicada pelo janelador de análise 1120 (por exemplo, razão = 4) será igual à razão da maior taxa efetiva de amostragem e à razão de amostragem original. Assim, um sinal de saída 1195 pode ser obtido e que tenha a mesma característica de sobreposição que o sinal original (com taxa de amostragem reduzida) 1115. O sinal de saída 1195 provido pelo aparelho 1100 pode ainda ser processado iniciando a partir do manipulador de reconstrução de alta frequência 130 como mostrado na Fig. 1a para finalmente obter um sinal replicado estendido em largura de banda.

Deve ser notado que na configuração da Fig. 11, todos os diferentes algoritmos de patching são feitos no mesmo domínio, por exemplo, no domínio da frequência. O domínio pode ser o domínio QMF, já que é feito em SBR ou em qualquer outro domínio, como o Fourier transposto. A geração real de dados de patch pode ser feita em um domínio diferente. Entretanto, nesse caso, todo o patching, é sempre feito no mesmo domínio.

Além disso, diferentes modelos fonte podem ser associados ao patching considerado na seleção. Por exemplo, um

modelo de origem de fala usado na extensão de largura de banda de fala, como descrito em *Frederik Nagel, Sascha Disch, "A harmonic bandwidth extension method for audio codecs," ICASSP International Conference on Acoustics, Speech and Signal Processing, IEEE CNF, Taipei, Taiwan, April 2009*, pode ser escolhido para sinais de fala, enquanto um modelo de origem estacionário pode ser adotado para música estacionária. Da mesma forma, como antes descrito, transientes podem ter seus modelos próprios para o patching.

Além disso, por meio de análise de sobreposição e de janelas de síntese para a transposição tempo-frequência, são garantidas transições suaves entre diferentes esquemas de patching. Alternativamente, podem ser usadas janelas especiais para análise e síntese para tornar possível a sobreposição mais baixa.

Em resumo, na configuração da Fig. 11, os métodos de patching podem ser selecionados entre uma simples operação de cópia das seções das frequências vizinhas, um esquema de transposição harmônica baseado em vocoder de fase, e um esquema de transposição harmônica baseado em vocoder de fase que inclua a cópia das seções das frequências vizinhas.

Apesar de a presente invenção ter sido descrita no contexto dos diagramas de blocos onde os blocos representam componentes reais ou lógicos de hardware, a presente invenção também pode ser implementada por um método computadorizado. Nesse último caso, os blocos representam as etapas correspondentes do método, onde essas etapas significam as funcionalidades realizadas pelos correspondentes blocos lógicos ou físicos de hardware.

As configurações descritas são meramente

ilustrativas dos princípios da presente invenção. Entende-se que modificações e alterações das disposições e dos detalhes aqui descritos ficarão aparentes para os técnicos no assunto. Pretende-se, portanto, que seja somente limitada pelo escopo das
5 reivindicações da patente pendente, e não pelos detalhes específicos apresentados como descrição e explicação das configurações apresentadas.

Dependente de determinadas exigências de implementação do método do invento, o método do invento pode ser
10 implementado em hardware ou em software. A implementação pode ser feita usando uma mídia de armazenagem digital, em particular um disco, um DVD ou um CD dotado de sinais de controle de leitura eletrônica, que cooperem com os sistemas de computador programável, de maneira que os métodos do invento sejam
15 realizados. Em geral, a presente invenção pode, portanto ser implementada como um produto de programa de computador, com um programa armazenado em um portador de leitura por máquina, o código do programa sendo operado para a realização dos métodos do invento quando o produto de programa de computador opera em um
20 computador. Em outras palavras, os métodos do invento são, portanto, um programa de computador tendo um código de programa para a realização de pelo menos um dos métodos do invento quando o programa de computador opera em um computador. O sinal de áudio codificado do invento pode ser armazenado em qualquer mídia de
25 armazenagem de leitura por máquina, como uma mídia de armazenagem digital.

As configurações da presente invenção permitem que a extensão de largura de banda leve em conta o som, o hardware

e características de sinal para o processo de patching. A decisão sobre o melhor patching que pode se adequar pode ser tomada dentro de loops abertos ou fechados. Portanto, a qualidade de restituição pode ser controlada e aperfeiçoada.

5 O conceito apresentado também tem a vantagem de poder ser facilmente alcançada uma transição suave entre os diferentes algoritmos de patching, permitindo uma adaptação rápida e precisa da extensão de largura de banda com base no sinal.

10 As aplicações mais importantes são os decodificadores de áudio, que são geralmente implementados em dispositivos manuais, operando assim com baterias.

REIVINDICAÇÕES

1. "APARELHO PARA A GERAÇÃO DE UM SINAL DE ÁUDIO DE SÍNTESE E PARA A CODIFICAÇÃO DE UM SINAL DE ÁUDIO", sendo o aparelho (100; 200; 700; 800; 1100) para a geração de um sinal de áudio de síntese (145) usando um sinal de controle de patching (119; 1155), o aparelho (100; 200; 700; 800; 1100) compreendendo:

um primeiro conversor (110; 1130) para a conversão de um primeiro ou segundo período de tempo (107-1; 107-2; 1125) de um sinal de áudio (105; 1101) em uma representação espectral (115; 1135-1);

um gerador de patch no domínio espectral (120; 1141) para a realização de uma pluralidade (117-1) de diferentes algoritmos de patching no domínio espectral, caracterizado por cada algoritmo de patching gerar uma representação espectral modificada (125; 1149) compreendendo componentes espectrais em uma banda superior de frequência (220) obtida a partir dos componentes espectrais correspondentes em uma banda de frequência de núcleo (210) do sinal de áudio (105; 1101), e onde o gerador de patch no domínio espectral (120; 1141) é configurado para selecionar um primeiro algoritmo de patching no domínio espectral (117-2) a partir da pluralidade (117-1) de algoritmos de patching para o primeiro período de tempo (107-1) e um segundo algoritmo de patching no domínio espectral (117-3) a partir da pluralidade (117-1) de algoritmos de patching para um segundo período de tempo (107-2) que é diferente do

primeiro período de tempo (107-1) de acordo com o sinal de controle de patching (119; 1155) para obter uma representação espectral modificada (125; 1149);

um manipulador de reconstrução de alta frequência (130) para a manipulação de uma representação espectral modificada (125; 1149) ou um sinal obtido a partir de uma representação espectral modificada (125; 1195) de acordo com um parâmetro de replicação de banda espectral (127) para obter um sinal ampliado de largura de banda (135); e

um combinador (140) para combinar o sinal de áudio (105; 1101) tendo componentes espectrais na banda de frequência de núcleo (210) ou um sinal obtido a partir do sinal de áudio (105; 1101) com o sinal ampliado de largura de banda (135) para obter o sinal de áudio de síntese (145),

onde o gerador de patch no domínio espectral (120) é configurado para realizar pelo menos dois diferentes algoritmos de patching no domínio espectral a partir de um grupo (203) de algoritmos de patching no domínio espectral, o grupo (203) de algoritmos de patching compreendendo:

um primeiro algoritmo de patching (205-1) compreendendo uma transposição harmônica baseada em um vocoder de fase simples e em funcionalidades de replicação de banda espectral de cópia não harmônica,

um segundo algoritmo de patching (205-2) compreendendo uma transposição harmônica baseada em um

vocoder de fase múltipla,

um terceiro algoritmo de patching (205-3) compreendendo funcionalidades de replicação de banda espectral de cópia não harmônica, e

um quarto algoritmo de patching (205-4) compreendendo uma distorção não linear

2. Aparelho (100; 200; 700; 800; 1100), de acordo com a reivindicação 1, caracterizado por o gerador de patch no domínio espectral (120; 1141) ser implementado para operar em um domínio espectral e não em um domínio do tempo.

3. Aparelho, de acordo com a reivindicação 1, caracterizado por o primeiro algoritmo de patching (205-1) compreender, como a transposição harmônica baseada em um vocoder de fase simples (305), o uso de um fator de extensão de largura de banda (σ) de dois controlando uma transformada de uma banda de frequência fonte (310) extraída da banda de frequência de núcleo (210) em uma primeira banda de frequência alvo (310'), onde as fases das componentes espectrais na banda de frequência fonte (310) são multiplicadas pelo fator de extensão de largura de banda (σ) de maneira que a primeira banda de frequência alvo (310) tenha frequências que variam da frequência de cruzamento (f_x) até o dobro da frequência de cruzamento (f_x), e

onde o primeiro algoritmo de patching (205-1) compreende, como as funcionalidades de replicação de banda espectral de cópia não harmônica (315), a

transformação dos componentes espectrais na primeira banda de frequência alvo (310') em uma segunda banda de frequência alvo (320') por meio de uma primeira cópia, de maneira que a segunda banda de frequência alvo (320') tenha frequências que variem do dobro da frequência de cruzamento (f_x) a três vezes a frequência de cruzamento (f_x), e a transformação dos componentes espectrais na segunda banda de frequência alvo (320') em uma terceira banda de frequência alvo (330') por meio de uma segunda cópia de maneira que a terceira banda de frequência alvo (330') tenha frequências que variem de três vezes da frequência de cruzamento (f_x) a quatro vezes da frequência de cruzamento (f_x) incluídas na banda superior de frequência (220), e

onde a banda superior de frequência (220) compreende a primeira banda de frequência alvo (310'), a segunda banda de frequência alvo (320') e a terceira banda de frequência alvo (330').

4. Aparelho, de acordo com a reivindicação 1, caracterizado por o segundo algoritmo de patching (205-2), o segundo algoritmo de patching (205-2) compreender, como a transposição harmônica com base em um vocoder de fase múltipla (405), o uso de um primeiro fator de extensão de largura de banda (σ_1) de dois controlando uma transformada a partir de uma primeira banda de frequência (410) extraída da banda de frequência de núcleo (210) em uma primeira banda de frequência alvo (410'), em que fases dos componentes espectrais na primeira banda de frequência fonte (410) são multiplicadas pelo primeiro fator de

extensão de largura de banda(σ_1) de maneira que a primeira banda de frequência alvo (410') tenha frequências que variem da frequência de cruzamento (f_x) até o dobro da frequência de cruzamento (f_x),

onde o segundo algoritmo de patching (205-2) compreende o uso de um segundo fator de extensão de largura de banda(σ_2) de três, controlando uma transformada de uma segunda banda de frequência fonte (420-1, 420-2) extraída da banda de frequência de núcleo (210) em uma segunda banda de frequência alvo (420', 420''), em que fases das componentes espectrais na segunda banda de frequência fonte (420-1, 420-2) são multiplicadas pelo segundo fator de extensão de largura de banda(σ_2) de maneira que a segunda banda de frequência alvo (420', 420'') tenha frequências que variem do dobro da frequência de cruzamento (f_x) a três vezes a frequência de cruzamento (f_x) ou variando entre a frequência de cruzamento (f_x) a três vezes a frequência de cruzamento (f_x),

onde o segundo algoritmo de patching (205-2) compreende o uso de um terceiro fator de extensão de largura de banda(σ_3) de quatro, controlando uma transformada de uma terceira banda de frequência fonte (430-1, 430-2) extraída da banda de frequência de núcleo (210) em uma terceira banda de frequência alvo (430', 430''), em que as fases dos componentes espectrais na terceira banda de frequência fonte (430', 430'') são multiplicadas pelo terceiro fator de extensão de largura de banda(σ_3) de maneira que a terceira banda de frequência

alvo (430', 430'') tenha frequências que variem de três vezes a frequência de cruzamento (f_x) a quatro vezes a frequência de cruzamento (f_x) ou variando entre a frequência de cruzamento (f_x) a quatro vezes a frequência de cruzamento (f_x) incluída na banda superior de frequência (220), e

onde a banda superior de frequência (220) compreende a primeira banda de frequência alvo (410'), a segunda banda de frequência (420', 420'') e a terceira banda de frequência alvo (430', 430'').

5. Aparelho, de acordo com a reivindicação 1, caracterizado por o terceiro algoritmo de patching (205-3) compreender, como as funcionalidades de replicação de banda espectral de cópia não harmônica (505),

a transformação dos componentes espectrais em uma banda de frequência fonte (510) sendo a banda de frequência de núcleo (210) em uma primeira banda de frequência alvo (510') por meio de uma primeira cópia de maneira que a primeira banda de frequência alvo (510') tenha frequências que variem da frequência de cruzamento (f_x) até o dobro da frequência de cruzamento (f_x),

a transformação dos componentes espectrais na primeira banda de frequência alvo (510') em uma segunda banda de frequência alvo (520') por meio de uma segunda cópia, de maneira que a segunda banda de frequência alvo (520') tenha frequências que variem do dobro da frequência de cruzamento (f_x) a três vezes a frequência de cruzamento (f_x), e

a transformação dos componentes espectrais na segunda banda de frequência alvo (520') em uma terceira banda de frequência alvo (530') por meio de uma terceira cópia, de maneira que a terceira banda de frequência alvo (530') tenha frequências que variem de três vezes a frequência de cruzamento (f_x) a quatro vezes a frequência de cruzamento (f_x) incluída na banda superior de frequência (220), e

onde a banda superior de frequência (220) compreende a primeira banda de frequência alvo (510'), a segunda banda de frequência alvo (520') e a terceira banda de frequência alvo (530').

6. Aparelho, de acordo com a reivindicação 1, caracterizado por o quarto algoritmo de patching (205-4) compreender, como a distorção não linear, a geração de componentes espectrais na banda superior de frequência (220) com frequências variando entre a frequência de cruzamento (f_x) até quatro vezes a frequência de cruzamento (f_x).

7. Aparelho (700), de acordo com uma das reivindicações anteriores, caracterizado por o aparelho (700) não compreendendo um transformador tempo/frequência (710) para a transformação de um sinal de domínio do tempo (705) obtida a partir da representação espectral modificada (125) no domínio espectral.

8. Aparelho (800), de acordo com uma das reivindicações anteriores, o aparelho (800) ainda compreendendo:

um segundo conversor (810) para a conversão de uma representação espectral modificada (125) no domínio do tempo, caracterizado por o segundo conversor (810) ser adaptado para adaptar uma síntese combinada a uma análise aplicada pelo primeiro conversor (110),

onde o primeiro conversor (110) é configurado para fazer uma conversão tendo um primeiro comprimento de conversão (111), e

onde o segundo conversor (810) é configurado para fazer uma conversão tendo um segundo comprimento de conversão, o segundo comprimento de conversão dependendo de uma característica da extensão de largura de banda em que são levados em conta uma razão da frequência máxima (f_{\max}) na banda superior de frequência (220) e a frequência de cruzamento (f_x) na banda de frequência de núcleo (210) e o primeiro comprimento de conversão (111).

9. Aparelho (900; 1000) para a codificação de um sinal de áudio (105) caracterizado por o sinal de áudio (105) compreender uma banda de frequência de núcleo (210) e uma banda superior de frequência (220), o aparelho (900; 1000) compreendendo:

um codificador núcleo (910) para a codificação do sinal de áudio (105) dentro da banda de frequência de núcleo (210);

um extrator paramétrico (920) para a extração de um sinal de controle de patching (119) do sinal de áudio (105), o sinal de controle de patching (119)

indicando um algoritmo de patching selecionado a partir da pluralidade (117-1) de diferentes algoritmos de patching no domínio espectral, o algoritmo de patching selecionado a ser realizado em um domínio espectral para a geração de um sinal de áudio de síntese em um decodificador de extensão de largura de banda; e

um calculador paramétrico (930) para o cálculo de um parâmetro de replicação de banda espectral (127) da banda superior de frequência (220).

10. Aparelho (1000) para a codificação de acordo com a reivindicação 9, caracterizado por o extrator paramétrico (920) ser configurado para determinar a partir da pluralidade (117-1) de diferentes algoritmos de patching no domínio espectral do algoritmo de patching selecionado, o algoritmo de patching selecionado baseando-se em uma comparação do sinal de áudio (105) ou de um sinal obtido a partir do sinal de áudio (105) com uma pluralidade (1005) de sinais ampliados de largura de banda tendo sido obtidos pela realização da pluralidade (117-1) de algoritmos de patching no domínio espectral e a manipulação de uma representação espectral modificada (125) de uma período de tempo do sinal de áudio (105).

11. Processo (100; 200; 700; 800; 1100) para a geração de um sinal de áudio de síntese (145) usando um sinal de controle de patching (190; 1155) em que (100; 200; 700; 800; 1100) compreender:

a conversão (110; 1130) de uma primeira ou segunda período de tempo (107-1; 107-2; 1125) de um sinal

de áudio (105; 1101) em uma representação espectral (115; 1135-1);

a realização de (120; 1141) uma pluralidade (117-1) de diferentes algoritmos de patching no domínio espectral, caracterizado por cada algoritmo de patching gerar uma representação espectral modificada (125; 1149) compreendendo componentes espectrais em uma banda superior de frequência (220) obtida a partir dos componentes espectrais correspondentes em uma banda de frequência de núcleo (210) do sinal de áudio (105; 1101), e selecionando (120; 1141) um primeiro algoritmo de patching no domínio espectral (117-2) a partir da pluralidade (117-1) de algoritmos de patching para um primeiro período de tempo (107-1) e um segundo algoritmo de patching no domínio espectral (117-3) a partir da pluralidade (117-1) de algoritmos de patching para um segundo período de tempo (107-2) que é diferente do primeiro período de tempo (107-1) de acordo com o sinal de controle de patching (119; 1155) para obter uma representação espectral modificada (125; 1149);

a manipulação (130) de uma representação espectral modificada (125; 1149) ou de um sinal obtido a partir de uma representação espectral modificada (125; 1195) de acordo com um parâmetro de replicação de banda espectral (127) para a obtenção de um sinal ampliado de largura de banda (135); e

a combinação (140) do sinal de áudio (105; 1101) tendo componentes espectrais na banda de frequência

de núcleo (210) ou de um sinal obtido a partir do sinal de áudio (105; 1101) com o sinal ampliado de largura de banda (135) para obter o sinal de áudio de síntese (145),

onde a realização (120; 1141) da pluralidade (117-1) de diferentes algoritmos de patching no domínio espectral compreende a realização de pelo menos dois diferentes algoritmos de patching no domínio espectral a partir de um grupo (203) de algoritmos de patching no domínio espectral, o grupo (203) de algoritmos de patching compreendendo:

um primeiro algoritmo de patching (205-1) compreendendo uma transposição harmônica baseada em um vocoder de fase simples e em funcionalidades de replicação de banda espectral de cópia não harmônica,

um segundo algoritmo de patching (205-2) compreendendo uma transposição harmônica baseada em um vocoder de fase múltipla,

um terceiro algoritmo de patching (205-3) compreendendo funcionalidades de replicação de banda espectral de cópia não harmônica, e

um quarto algoritmo de patching (205-4) compreendendo uma distorção não linear

12. Processo (900; 1000) para a codificação de um sinal de áudio (105), caracterizado por o sinal de áudio (105) compreender uma banda de frequência de núcleo (210) e uma banda superior de frequência (220), o método (900; 1000) compreendendo:

a codificação (910) do sinal de áudio (105)

dentro da banda de frequência de núcleo (210);

a extração (920) de um sinal de controle de patching (119) do sinal de áudio (105), o sinal de controle de patching (119) indicando um algoritmo de patching selecionado a partir da pluralidade (117-1) de diferentes algoritmos de patching no domínio espectral, o algoritmo de patching selecionado para ser realizado em um domínio espectral para a geração de um sinal de áudio de síntese em um decodificador de extensão de largura de banda; e

o cálculo (930) de um parâmetro de replicação de banda espectral (127) da banda superior de frequência (220).

13. Sinal de áudio codificado (935) caracterizado por compreender:

um sinal de áudio codificado (915) codificado no interior de uma banda de frequência de núcleo (210);

um sinal de controle de patching (119), o sinal de controle de patching (119) indicando um algoritmo de patching selecionado a partir da pluralidade (117-1) de diferentes algoritmos de patching no domínio espectral, o algoritmo de patching selecionado a ser realizado em um domínio espectral para a geração de um sinal de áudio de síntese (145) em um decodificador de extensão de largura de banda; e

um parâmetro de replicação de banda espectral (127) calculado a partir de uma banda superior de frequência (220) do sinal de áudio (105).

100

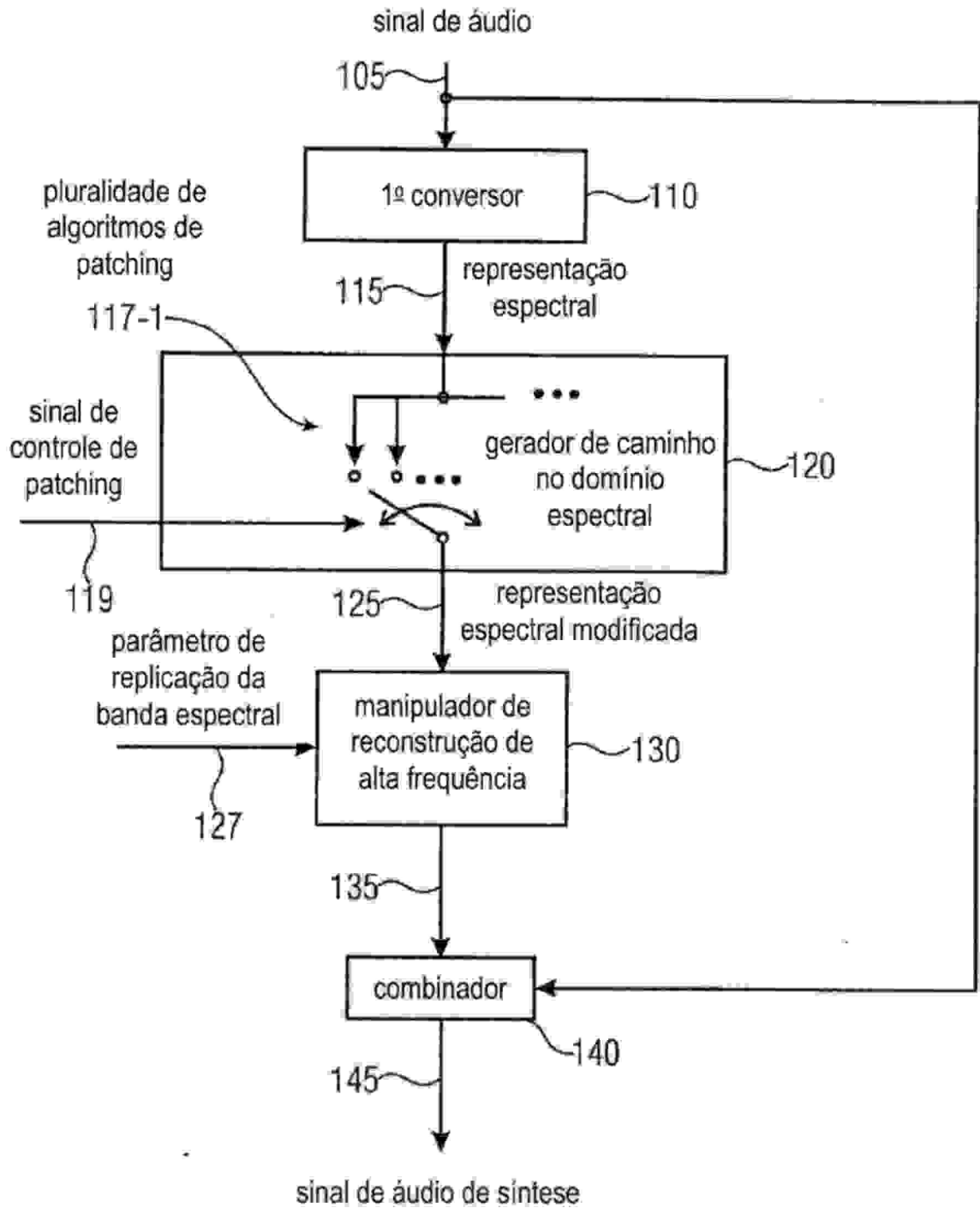


Figura 1A
(SINTETIZADOR DE SINAL DE ÁUDIO)

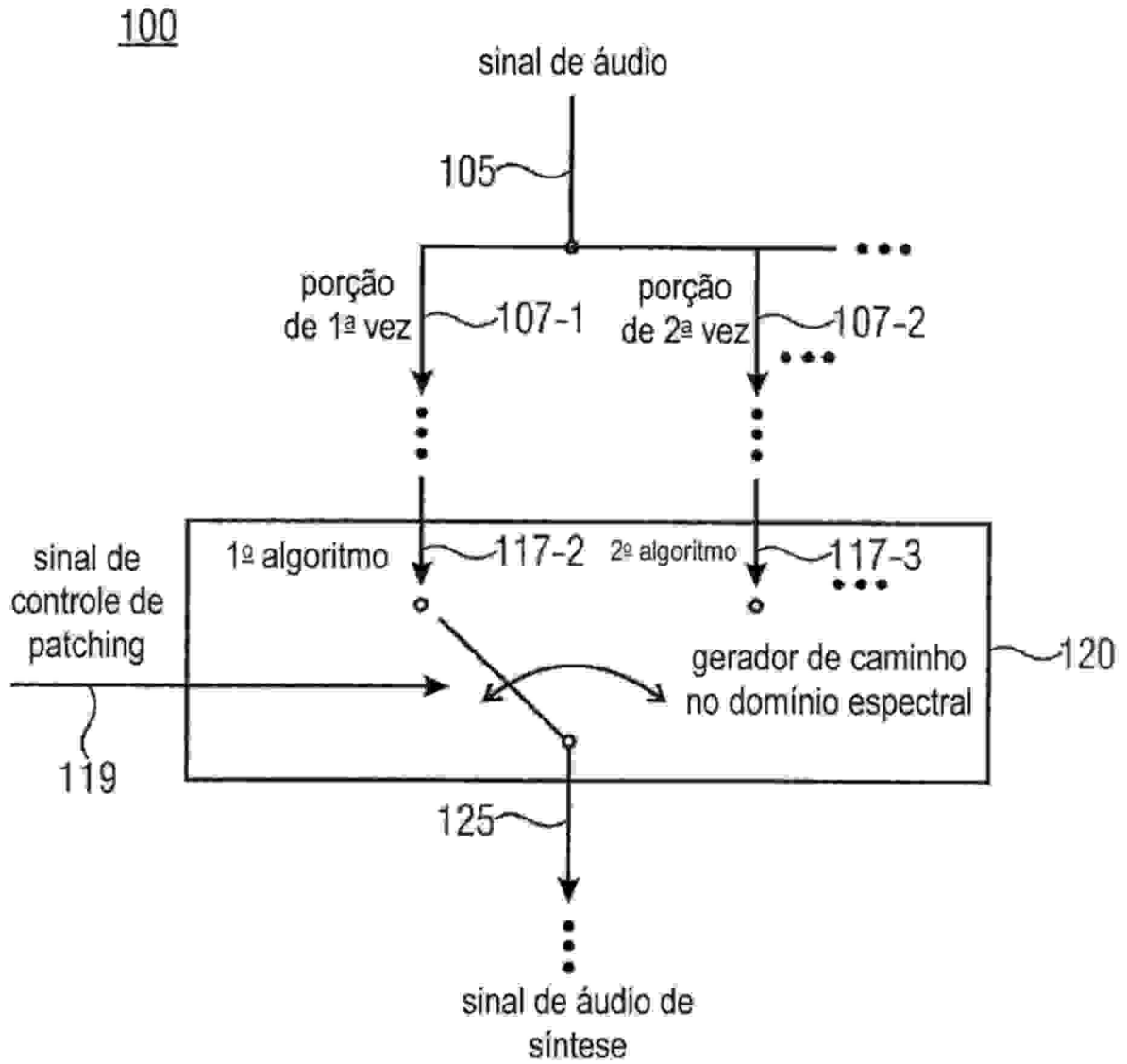


Figura 1B
(GERADOR DE PATCH NO DOMÍNIO ESPECTRAL)

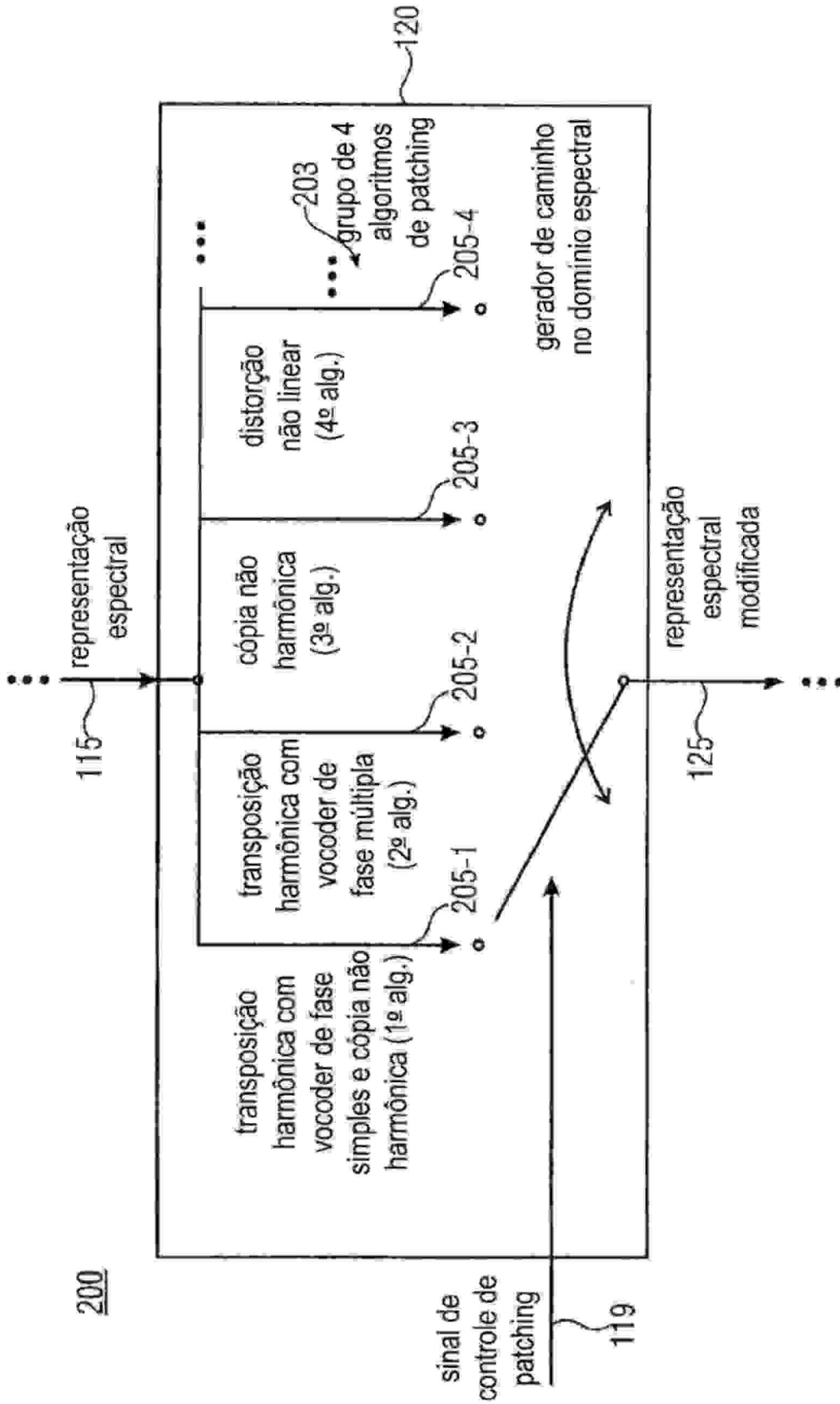


Figura 2A

200

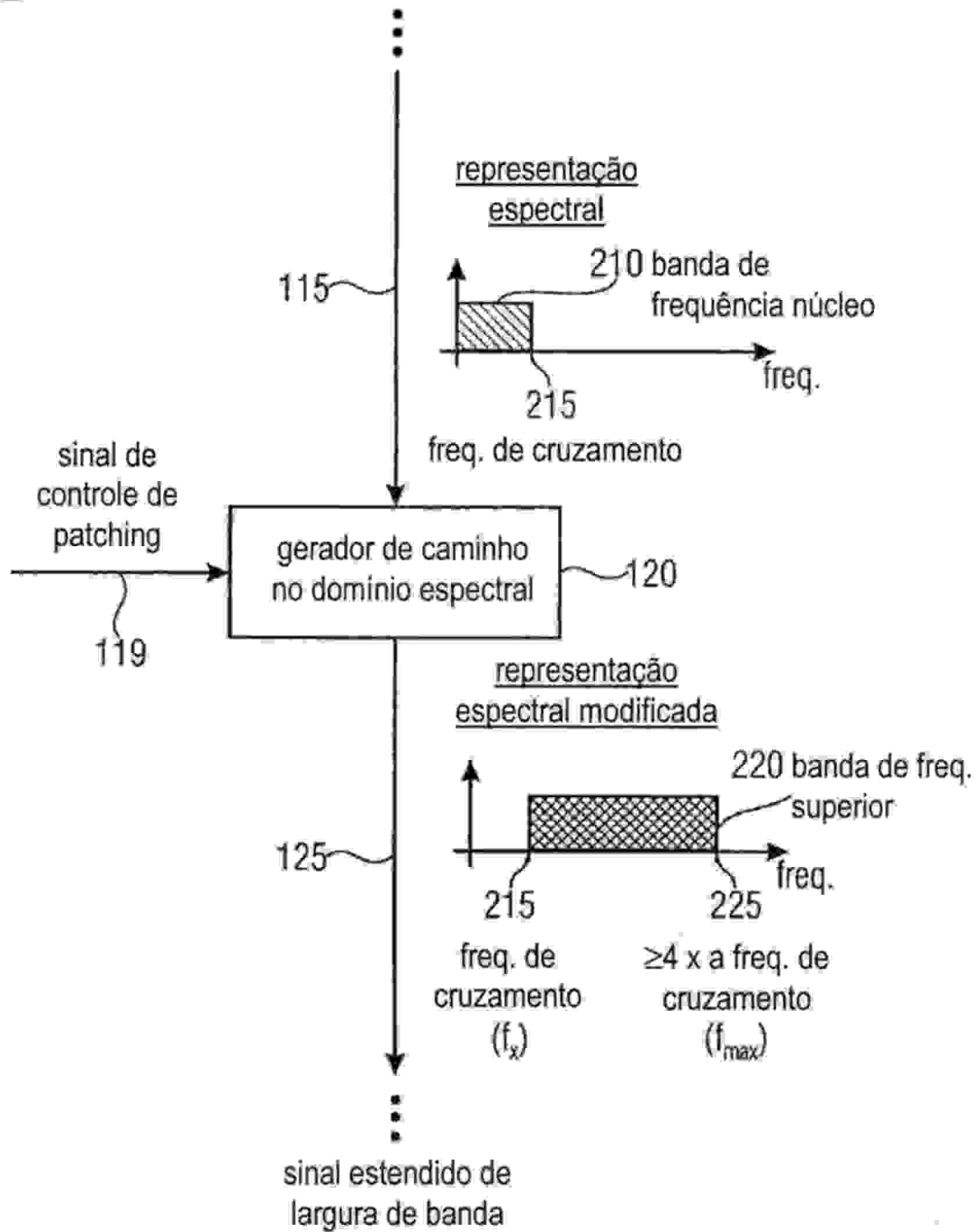


Figura 2B
(ESQUEMA DE EXTENSÃO DE LARGURA DE BANDA)

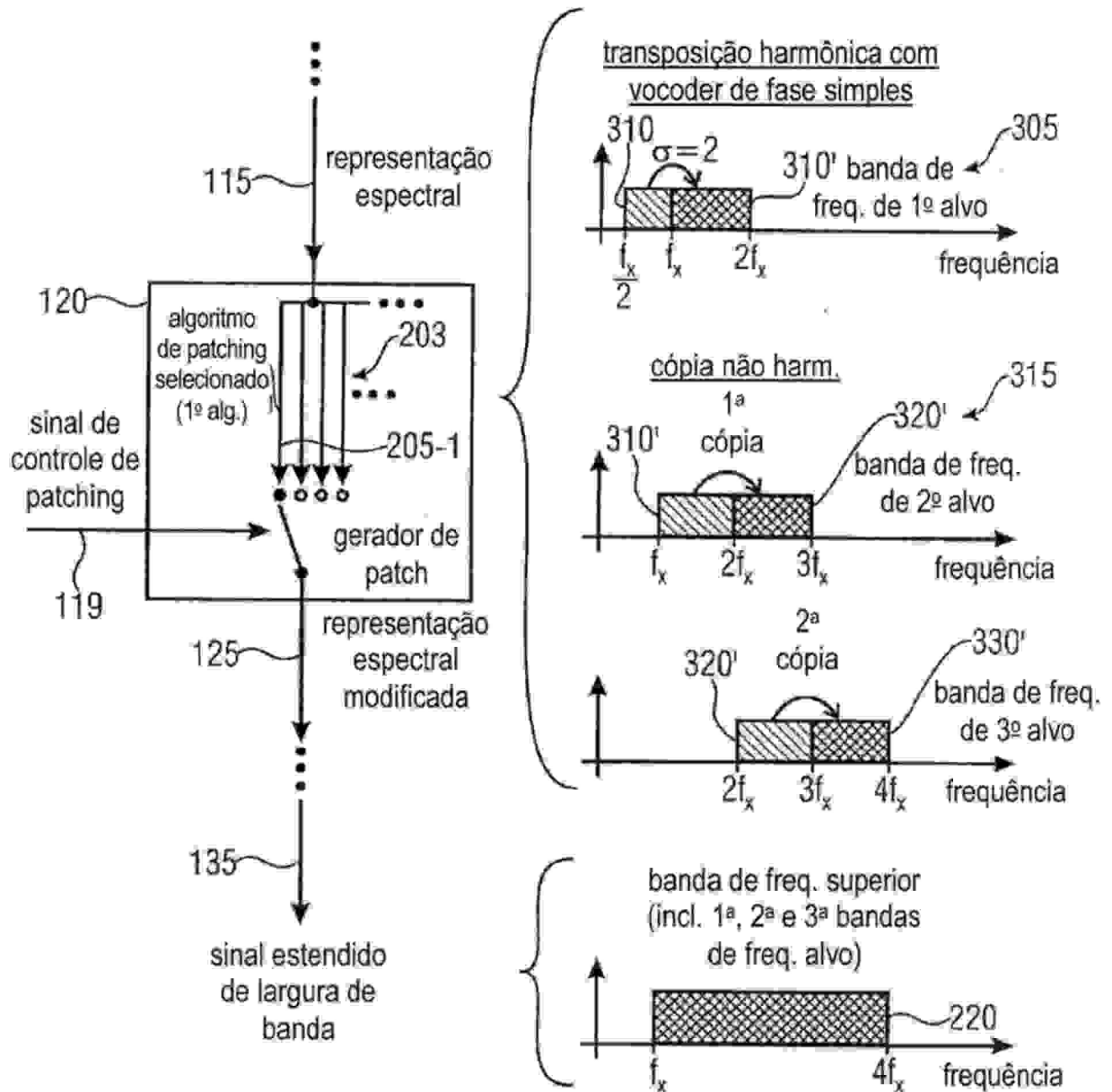


Figura 3
(PRIMEIRO ALGORITMO DE PATCHING)

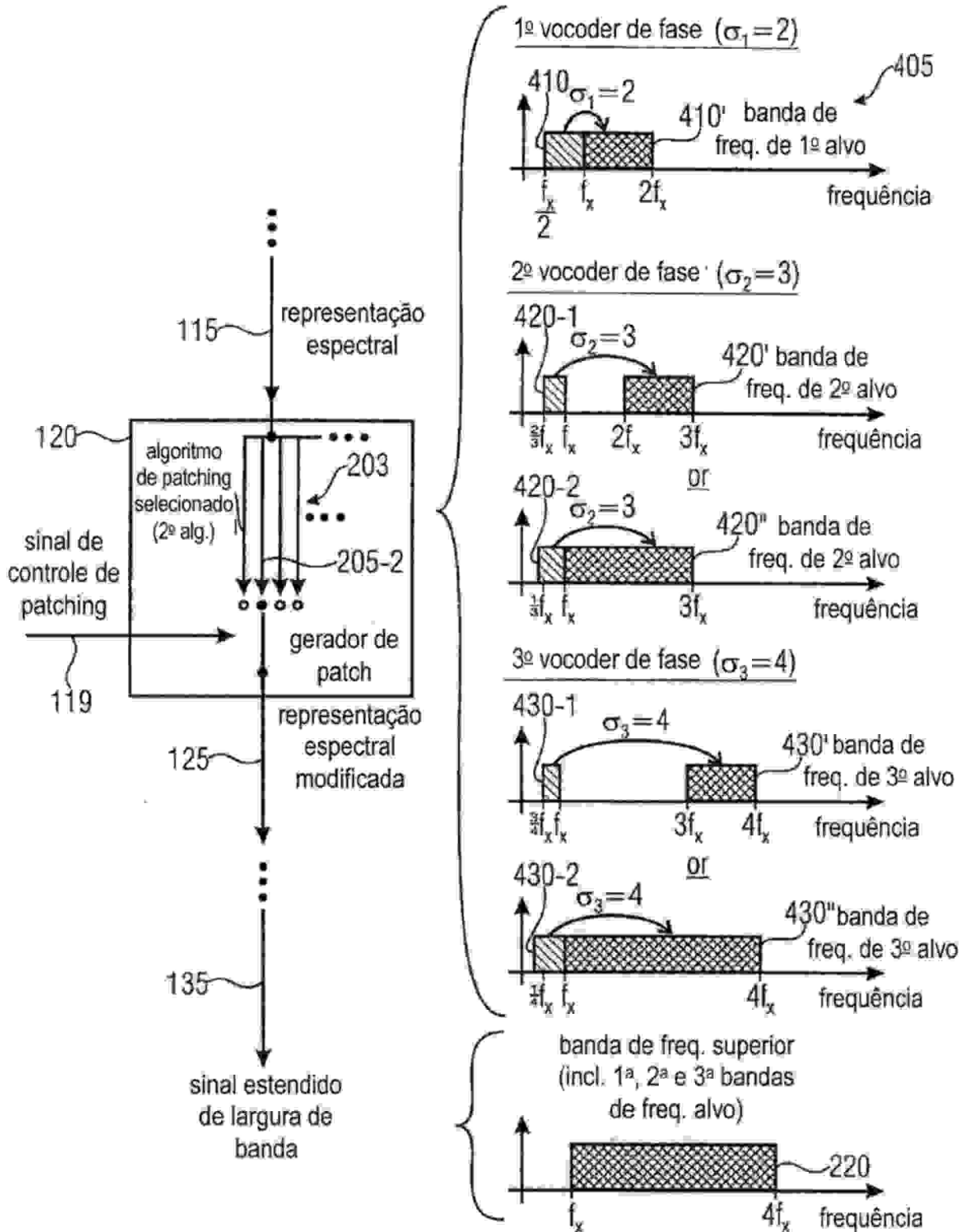


Figura 4
(SEGUNDO ALGORITMO DE PATCHING)

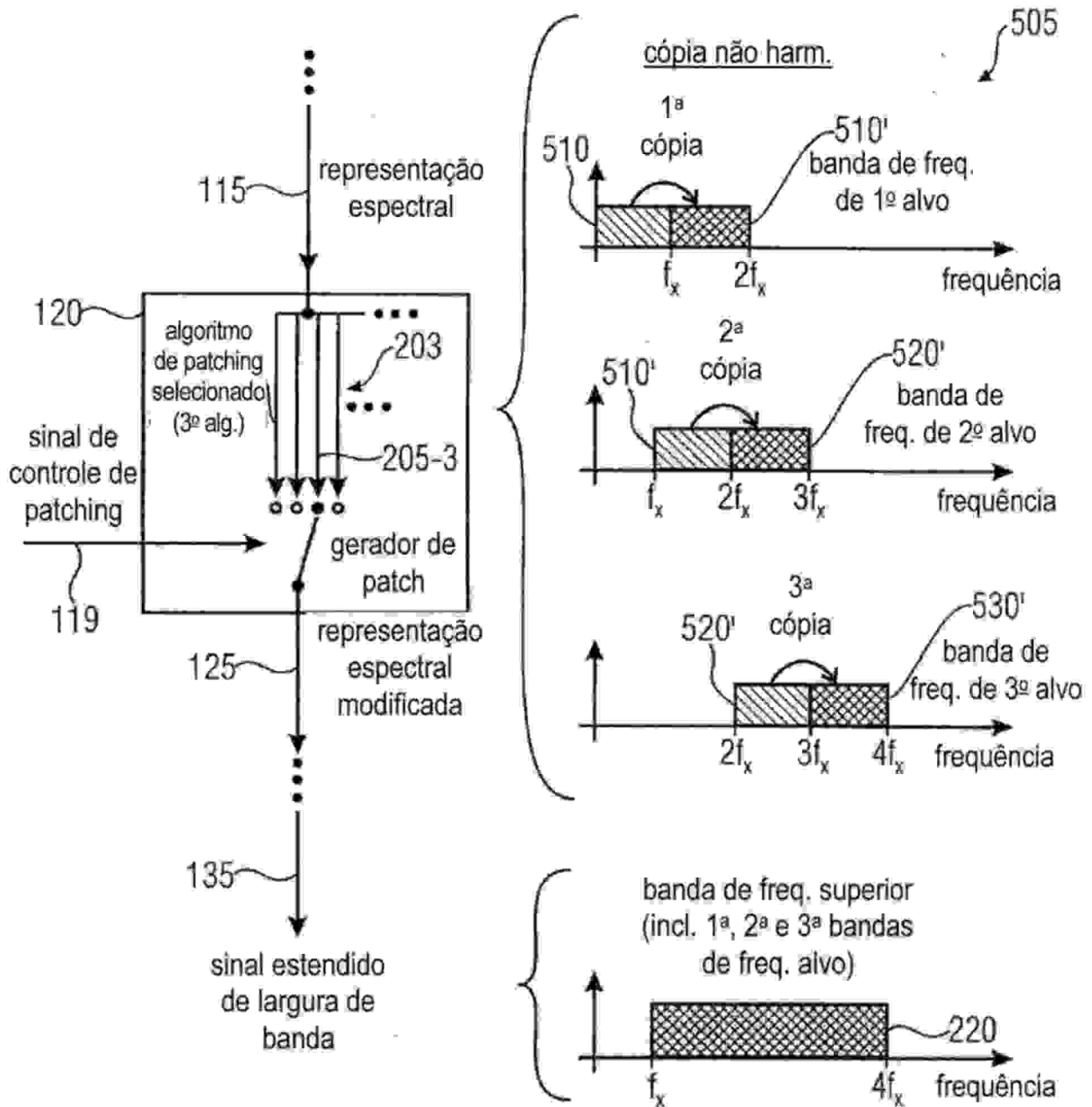


Figura 5
(TERCEIRO ALGORITMO DE PATCHING)

8/13

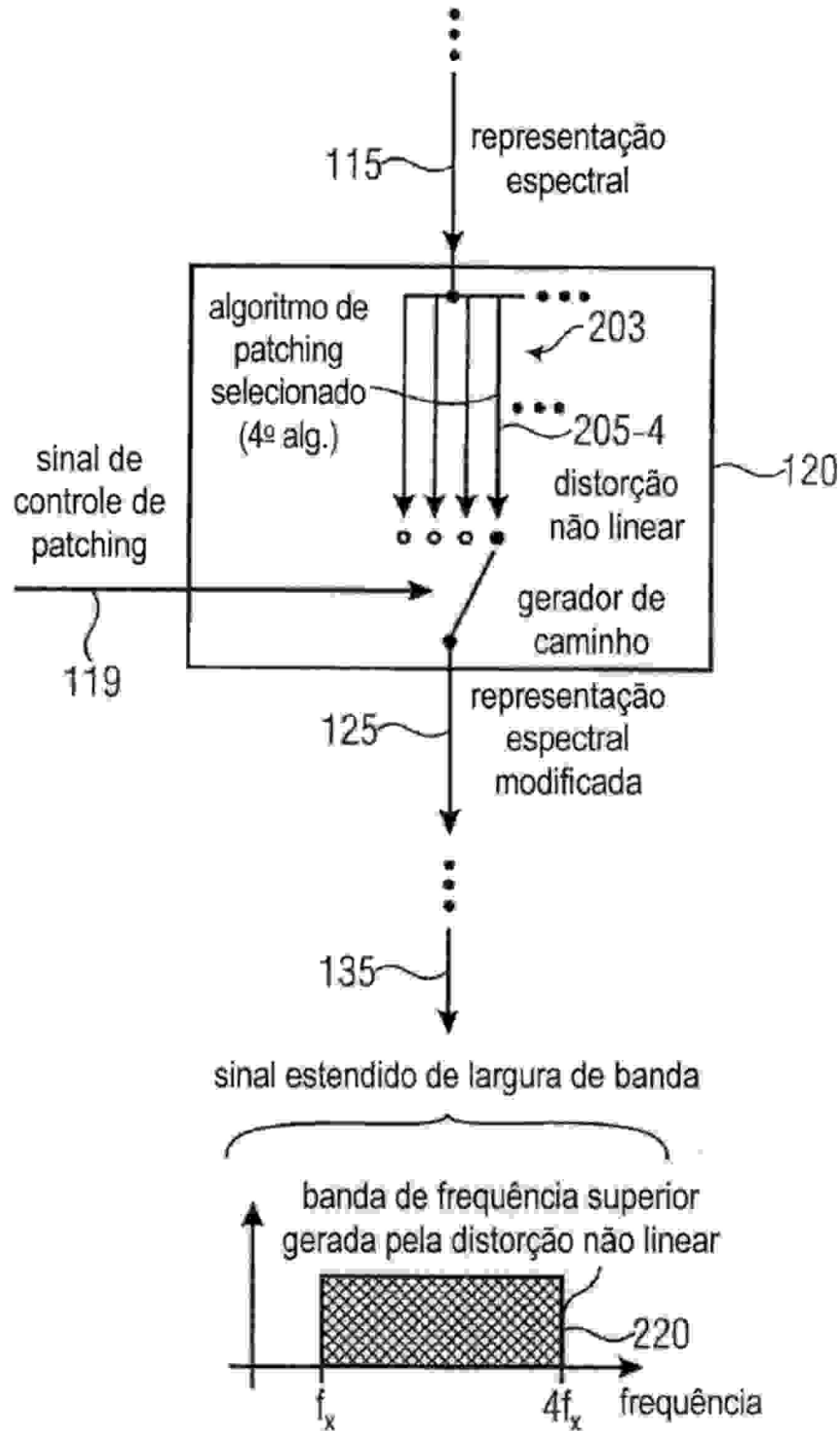
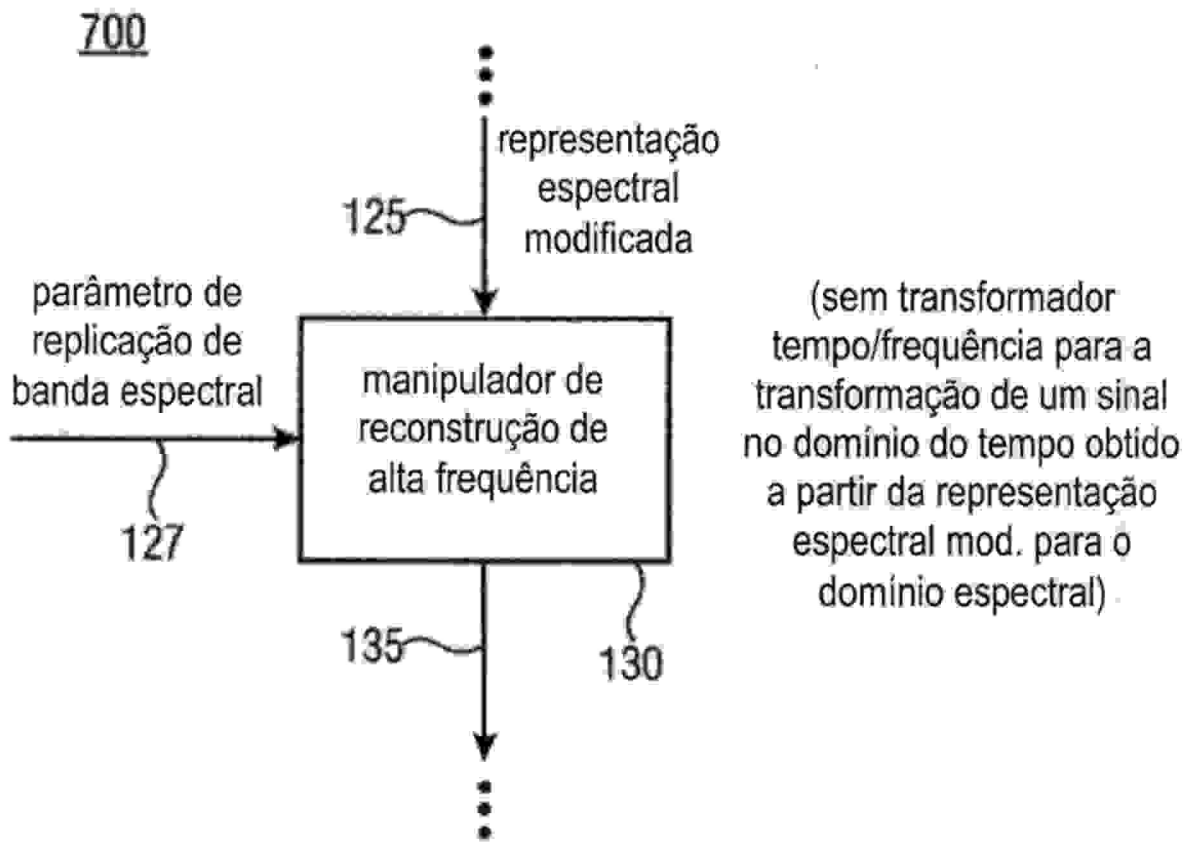


Figura 6
(QUARTO ALGORITMO DE PATCHING)

**Figura 7**

800

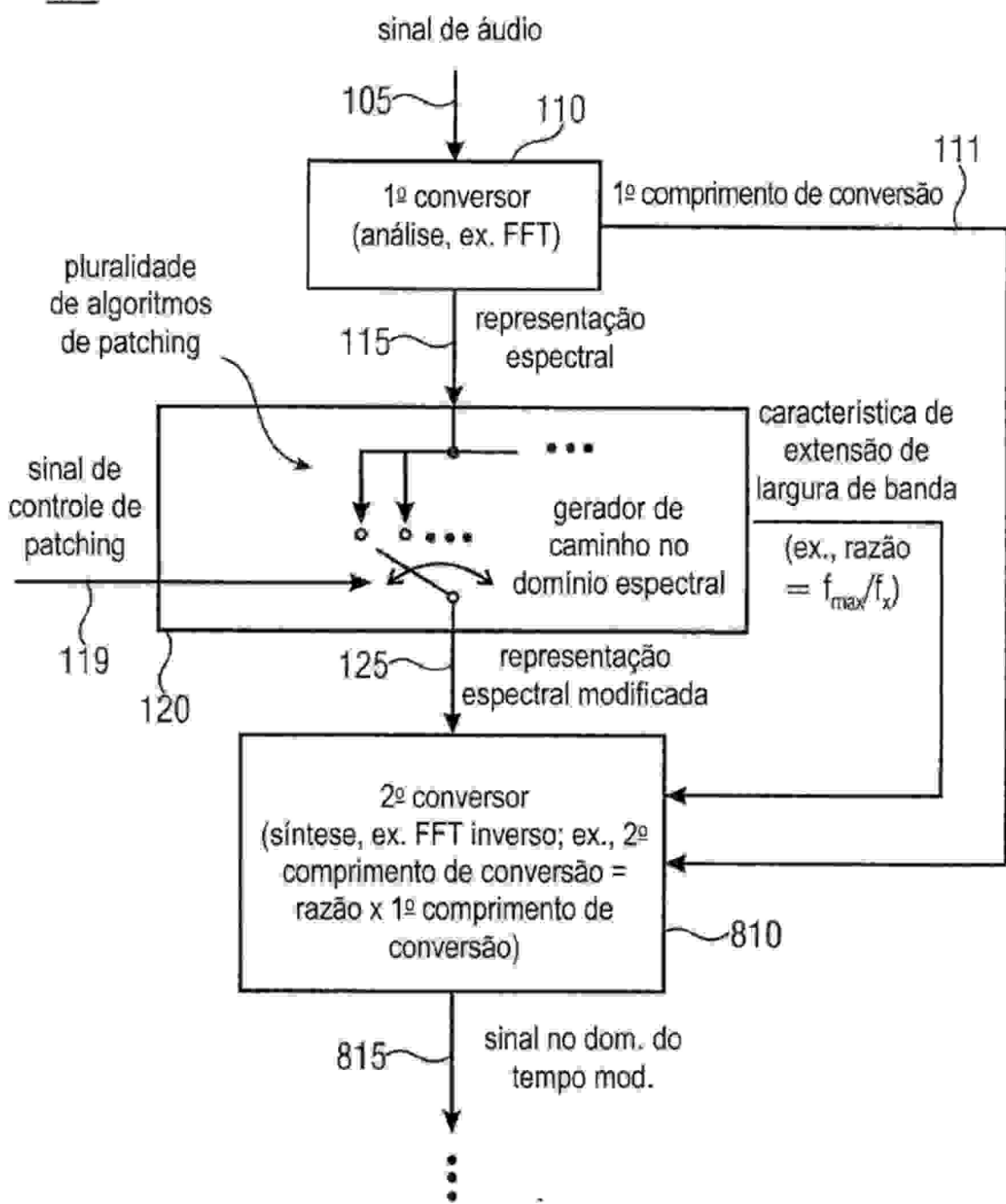


Figura 8
(SINTETIZADOR DE SINAL DE ÁUDIO)

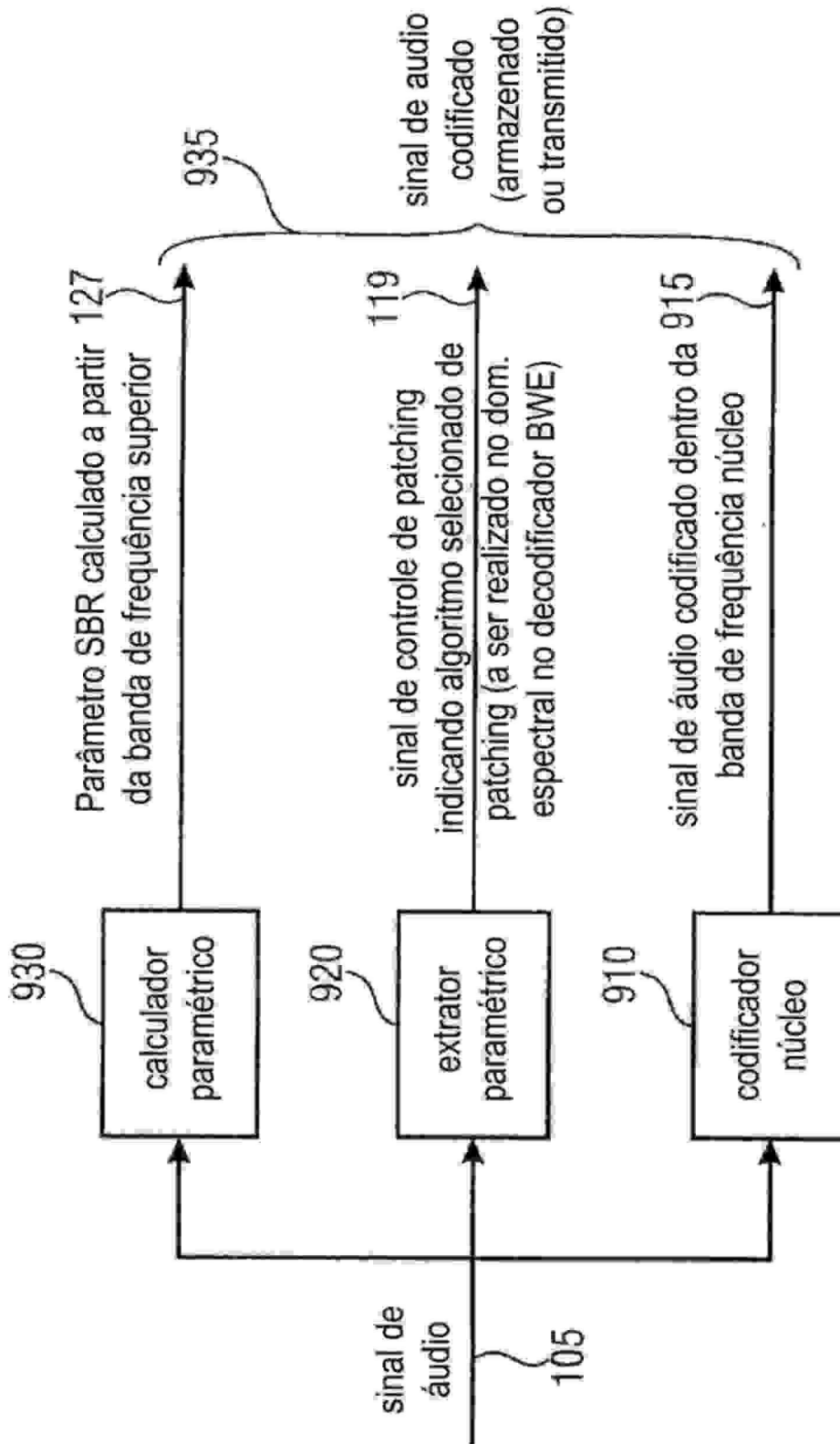


Figura 9
(CODIFICADOR)

Figura 11
(ESQUEMA DE PATCHING NO
DOMÍNIO DA FREQUÊNCIA)

