



INPI
Assinado
Digitalmente

REPÚBLICA FEDERATIVA DO BRASIL
MINISTÉRIO DA ECONOMIA
INSTITUTO NACIONAL DA PROPRIEDADE INDUSTRIAL

CARTA PATENTE Nº PI 0610441-0

O INSTITUTO NACIONAL DA PROPRIEDADE INDUSTRIAL concede a presente PATENTE DE INVENÇÃO, que outorga ao seu titular a propriedade da invenção caracterizada neste título, em todo o território nacional, garantindo os direitos dela decorrentes, previstos na legislação em vigor.

(21) Número do Depósito: PI 0610441-0

(22) Data do Depósito: 23/03/2006

(43) Data da Publicação do Pedido: 26/10/2006

(51) Classificação Internacional: H04S 7/00; G10L 19/02.

(30) Prioridade Unionista: US 60/671,381 de 13/04/2005.

(54) Título: MEDIÇÃO ECONÔMICA DE INTENSIDADE DE ÁUDIO CODIFICADO

(73) Titular: DOLBY LABORATORIES LICENSING CORPORATION, Sociedade Norte-Americana. Endereço: 100 POTRERO AVENUE, SAN FRANCISCO CA 94103, ESTADOS UNIDOS DA AMÉRICA(US)

(72) Inventor: BRETT GRAHAM CROCKETT; MICHAEL JOHN SMITHERS; ALAN JEFFREY SEEFELDT.

Prazo de Validade: 10 (dez) anos contados a partir de 02/01/2019, observadas as condições legais

Expedida em: 02/01/2019

Assinado digitalmente por:
Liane Elizabeth Caldeira Lage
Diretora de Patentes, Programas de Computador e Topografias de Circuitos Integrados

Relatório Descritivo da Patente de Invenção para "**MEDIÇÃO
ECONÔMICA DE INTENSIDADE DE ÁUDIO CODIFICADO**".

Campo Técnico

5 A presente invenção refere-se ao processamento de sinal de áudio. Mais particularmente, refere-se a um cálculo econômico de uma medição objetiva de intensidade de áudio codificado de baixa taxa de bits, tal como áudio codificado usando Dolby Digital (AC-3), Dolby Digital Plus ou Dolby E. "Dolby", "Dolby Digital", Dolby Digital Plus e "Dolby E" são marcas de indústria e comércio de Dolby Laboratories Licensing Corporation. Aspectos da invenção também podem ser utilizáveis com outros tipos de codificação de áudio.

Antecedentes da Técnica

Detalhes da codificação Dolby Digital são apresentados nas referências a seguir:

15 *ATSC Standard A52/A: Digital Audio Compression Standard (AC-3), Revision A*, Advanced Television Systems Committee, 20 de agosto de 2001. O documento A/52A está disponível na World Wide Web (WWW - Rede Mundial), em <http://www.atsc.org/standards.html>.

20 Flexible Perceptual Coding for Audio Transmission and Storage, de Craig C. Todd, e outros, 96th Convention of the Audio Engineering Society, 26 de fevereiro de 1994, Pré-impressão 3796;

"Design and Implementation of AC-3 Coders", de Steve Vernon, *IEEE Trans. Consumer Electronics*, Vol. 41, Nº 3, agosto de 1995.

25 "The AC-3 Multichannel Coder" de Mark Davis, Audio Engineering Society, Pré-impressão 3774, 95th AES Convention, outubro de 1993.

"High Quality, Low-Rate Audio Transform Coding for Transmission and Multimedia Applications", de Bosi e outros, Audio Engineering Society, Pré-impressão 3365, 93rd AES Convention, outubro de 1992.

30 Patentes norte-americanas 5.583.962; 5.632.005; 5.663.981; 5.727.119; 5.909.664; e 6.021.386.

Detalhes da codificação Dolby Digital Plus são apresentados em "Introduction to Dolby Digital Plus, na Enhancement to the Dolby Digital Co-

ding System", AES Convention Paper 6196, 117th AES Convention, 28 de outubro de 2004.

Detalhes da codificação Dolby E são apresentados em "*Efficient Bit Allocation, Quantization, and Coding in na Audio Distribution System*",
5 AES Pré-impressão 5068, 107th AES Conference, agosto de 1999 e "*Professional Audio Coder Optimized for Use with Video*", AES Pré-impressão 5033, 107th AES Conference, agosto de 1999.

10 Uma visão geral de vários codificadores perceptuais, incluindo codificadores Dolby, codificadores MPEG e outros, é apresentada em "*Overview of MPEG Audio: Current and Future Standards for Low-Bit-Rate Audio Coding*", de Karlheinz Brandenburg e Marina Bosi, J. Audio Eng. Soc., Vol. 45, Nº 1/2, janeiro/ fevereiro, 1997.

Existem muitos métodos para medir, objetivamente, a intensidade percebida de sinais de áudio. Exemplos de métodos incluem medidas de
15 potência ponderada (tais como LeqA, LeqB, LeqC), bem como medidas de intensidade baseadas em psicoacústica, tais como "*Acoustics - Method for Calculating Loudness Level*", ISO 532 (1975). Medidas de intensidade de potência ponderada processam o sinal de áudio de entrada através da aplicação de um filtro predeterminado que enfatiza frequências sensíveis, mais perceptivelmente, enquanto tira a ênfase de frequências sensíveis menos perceptivelmente e, então, determina a média da potência do sinal filtrado
20 através de uma extensão de tempo predeterminada. Métodos psicoacústicos, tipicamente, são mais complexos e objetivam modelar melhor o funcionamento do ouvido humano. Isso é obtido pela divisão do sinal de áudio em
25 bandas de frequência que imitam a resposta de frequência e a sensibilidade do ouvido e, então, manipulando e integrando essas bandas, enquanto se leva em conta o fenômeno psicoacústico, tal como mascaramento de frequência e temporal, bem como a percepção não linear de intensidade com intensidade de sinal variável. O objetivo de todos os métodos objetivos de
30 medição de intensidade é derivar uma medição numérica da intensidade que combina estreitamente com a percepção subjetiva da intensidade de um sinal de áudio.

A codificação perceptual ou codificação de áudio de baixa taxa de bits é usada, comumente para comprimir dados de sinais de áudio para armazenamento, transmissão e distribuição eficientes em aplicações, tais como televisão digital difundida e a sala de música *on-line* da Internet. A codificação perceptual obtém sua eficiência por meio da transformação do sinal de áudio em um espaço de informação, onde componentes de sinais e redundâncias, que são mascarados psicoacusticamente, podem ser facilmente descartados. A informação restante é compactada em um fluxo ou arquivo de informação digital. Tipicamente, a medição da intensidade do áudio representado pelo áudio codificado de baixa taxa de bits requer decodificação da recuperação de áudio no domínio de tempo (por exemplo, PCM), que pode ser intensiva, computacionalmente. Contudo, alguns sinais codificados perceptualmente de baixa taxa de bits contêm informação que pode ser útil a um método de medição de intensidade, assim, economizando o custo computacionalmente de decodificação completa do áudio. Dolby Digital (AC-3), Dolby Digital Plus e Dolby E estão entre esses sistemas de codificação de áudio.

Os codificadores de áudio perceptuais de baixa taxa de bits Dolby Digital, Dolby Digital Plus e Dolby E dividem sinais de áudio em segmentos de tempo de sobreposição, de janela (ou blocos de codificação de áudio) que são transformados em uma representação de domínio de frequências. A representação de domínio de frequências de coeficientes espectrais é expressa por uma notação exponencial, compreendendo conjuntos de um expoente e mantissas associadas. Os expoentes, que funcionam na maneira de fatores de escala, são compactados no fluxo de áudio codificado. As mantissas representam os coeficientes espectrais após eles terem sido normalizados pelos expoentes. Os expoentes são, então, passados através de um modelo perceptual de audição e usados para quantificar e compactar as mantissas no fluxo de áudio codificado. Com a decodificação, os expoentes são descompactados do fluxo de áudio codificado e, então, passados através do mesmo modelo perceptual para determinar como descompactar as mantissas. As mantissas são, então, descompactadas, combinadas com

os expoentes para criar uma representação de domínio de freqüências do áudio, que é, então, decodificada e convertida de volta a uma representação de domínio de tempo.

5 Como muitas medições de intensidade incluem cálculos de potências e espectro de potências, economias computacionais podem ser obtidas pela decodificação apenas parcial do áudio codificado e de baixa taxa de bits e passando a informação parcialmente decodificada (tal como o espectro de potência) para a medição de intensidade. A invenção é útil quando há uma necessidade de medir a intensidade, mas não de decodificar o áudio. Ela explora o fato de que uma medição de intensidade pode fazer uso de uma versão aproximada do áudio, essa aproximação, usualmente, não sendo adequada para audição. Um aspecto da presente invenção é o reconhecimento que uma representação bruta do áudio, que está disponível sem decodificar um fluxo de bits em muitos sistemas de codificação de áudio, 15 pode proporcionar uma aproximação do espectro de áudio que é utilizável na medição da intensidade do áudio. Na codificação de áudio Dolby Digital, Dolby Digital Plus e Dolby E, expoentes proporcionam uma aproximação do espectro de potência do áudio. Similarmente, em certos outros sistemas de codificação, fatores de escala, envelopes espectrais e coeficientes preditivos lineares podem proporcionar uma aproximação do espectro de potência do áudio. Esses e outros aspectos e vantagens da invenção serão melhor compreendidos à medida que o sumário e a descrição seguintes da invenção são lidos e compreendidos. 20

O documento US 2001/0027393 descreve um sistema de áudio 25 conferência constituído de N terminais conectados respectivamente a uma unidade de controle multiponto. Cada terminal é constituído de um codificador cuja entrada recebe dados de áudio para transmitir a outros terminais e cuja saída é conectada a uma entrada da unidade de controle multiponto. Cada terminal também tem um decodificador cuja entrada é conectada à 30 saída da unidade de controle multiponto e cuja saída entrega dados que são transmitidos ao terminal considerado pelos outros terminais. A unidade de controle multiponto é essencialmente constituída de um combinador que

combina os sinais presentes em suas entradas e entrega à entrada do decodificador de um terminal um sinal representativo da soma dos sinais entregues respectivamente por todos os decodificadores dos N terminais, exceto para o sinal vindo daquele terminal. A unidade de controle multiponto também tem N decodificadores parciais cujo objetivo é receber, respectivamente, os quadros de áudio produzidos pelos N terminais para decodificá-los e assim entregá-los às entradas do combinador. A unidade de controle multiponto tem N recodificadores parciais tendo saídas respectivamente conectadas às entradas dos decodificadores dos terminais e tendo entradas conectadas às saídas do combinador. O documento descreve calcular a energia total em cada banda de frequência.

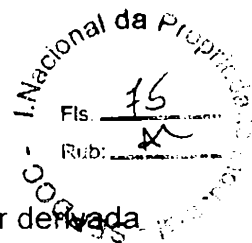
É um objetivo da presente invenção proporcionar uma medição computacionalmente econômica da intensidade percebida de áudio codificado de baixa taxa de bits.

Este objetivo é obtido pelo método de acordo com a reivindicação 1, As modalidades preferidas da invenção estão definidas nas reivindicações dependentes

Por conseguinte, o objetivo é obtido apenas decodificando parcialmente o material de áudio e passando a informação parcialmente decodificada à medição de intensidade de áudio. O método tira vantagem de propriedades específicas da informação de áudio parcialmente decodificada, tais como os expoentes em codificação de áudio Dolby Digital, Dolby Digital Plus e Dolby E.

Um primeiro aspecto da invenção mede a intensidade de áudio codificado em um fluxo de bits que inclui dados dos quais uma aproximação do espectro de potência do áudio pode ser derivada sem decodificar completamente o áudio através da derivação da aproximação do espectro de potência do áudio do fluxo de bits sem decodificar completamente o áudio e determinando uma intensidade aproximada do áudio em resposta à aproximação do espectro de potência do áudio.

Em outro aspecto da invenção, os dados podem incluir representações brutas do áudio e representações mais finas associadas do áudio, em



cujo caso, a aproximação do espectro de potência do áudio pode ser derivada das representações brutas do áudio.

Em um outro aspecto da invenção, o áudio codificado em um fluxo de bits pode ser áudio codificado de sub-bandas, cada sub-banda tendo um fator de escala e dados de amostra com ela associados e em que as representações brutas do áudio compreendem fatores de escala e as representações mais finas associadas do áudio compreendem dados de amostra associados com cada fator de escala.

Ainda em um outro aspecto da invenção, o fator de escala e os dados de amostra de cada sub-banda podem representar coeficientes espectrais na sub-banda por meio de notação exponencial em que o fator de escala compreende um expoente e os dados de amostra associados compreendem mantissas.

Ainda em um outro aspecto da invenção, o áudio codificado em um fluxo de bits pode ser áudio codificado preditivo linear em que as representações brutas do áudio compreendem coeficientes preditivos lineares e as representações mais finas do áudio compreendem informação de excitação associada com os coeficientes preditivos lineares.

Ainda em outro aspecto da invenção, as representações brutas do áudio podem compreender pelo menos um envelope espectral e as representações mais finas do áudio podem compreender componentes espectrais associados com o pelo menos um envelope espectral.

Ainda em outro aspecto da invenção, a determinação de uma intensidade aproximada do áudio em resposta à aproximação do espectro de potência do áudio pode incluir a aplicação de uma medição de intensidade de potência ponderada. A medição de intensidade de potência ponderada pode empregar um filtro que tira a ênfase de menos frequências perceptíveis e estabelece a média da potência do áudio filtrado através do tempo.

Ainda em outro aspecto da invenção, a determinação de uma intensidade aproximada do áudio em resposta à aproximação do espectro de potência do áudio pode incluir a aplicação de uma medição de intensidade psicoacústica. A medição de intensidade psicoacústica pode empregar um

modelo do ouvido humano para determinar intensidade específica em cada uma de uma pluralidade de bandas de frequências similares às bandas críticas do ouvido humano. Em um ambiente de codificador de sub-banda, as sub-bandas podem ser similares às bandas críticas do ouvido humano e a

5 medição de intensidade psicoacústica pode empregar um modelo do ouvido humano para determinar intensidade específica em cada uma das sub-bandas.

Aspectos da invenção incluem métodos que põem em prática as funções acima, meios para pôr em prática as funções, aparelho para colocar

10 em prática os métodos e um programa de computador, armazenado em um meio legível em computador para fazer com que um computador realize os métodos que põem em prática as funções acima.

Descrição dos Desenhos

A figura 1 mostra um diagrama em blocos funcional esquemático de uma disposição geral para medição da intensidade de áudio codificado de

15 baixa taxa de bits.

A figura 2 mostra um diagrama em blocos funcional esquemático generalizado de um decodificar Dolby Digital, Dolby Digital Plus e Dolby E.

As figuras 3a e 3b mostram diagramas em blocos funcionais esquemáticos de duas disposições gerais para calcular uma medição de volume/intensidade de som usando potência ponderada e medições baseadas em

20 psicoacústico, respectivamente.

A figura 4 mostra ponderações de frequência comuns usadas quando das medições da intensidade de som de acordo com a disposição do

25 exemplo da figura 3a.

Figura 5 é um diagrama de blocos funcional esquemático mostrando uma disposição geral mais econômica para medição da intensidade de som de áudio codificado de acordo com os aspectos da invenção.

As figuras 6a e 6b são diagramas de blocos funcionais esquemáticos da disposição mais econômica para medir a intensidade, incorporando as disposições de intensidade mostradas nos exemplos das figuras 3a

30 e 3b de acordo com os aspectos da invenção.

Melhor Modo para Realização da Invenção

Um benefício de aspectos da presente invenção é a medição da intensidade de áudio codificado de baixa taxa de bits, sem a necessidade de decodificar completamente o áudio para PCM, decodificação que inclui etapas onerosas de processamento de decodificação, tais como alocação de bits, dequantização, uma transformação inversa, etc. Aspectos da invenção reduzem grandemente as exigências de processamento sobrecarga computacional). Essa abordagem é benéfica, quando uma medição de intensidade é desejada, mas o áudio decodificado não é necessário.

Aspectos da presente invenção são utilizáveis, por exemplo, em ambientes, tais como, descritos em (1) Pedido de Patente Não-Provisório, pendente, dos Estados Unidos N.S. 10/884.117, depositado em 1 de julho de 2004, intitulado "*Method for Correcting Metadata Affecting the Playback Loudness and Dynamic Range of Audio Information*", de Smithers e outros; (2) Pedido de Patente Provisório Norte-americano N.S. 60/xxx,xxx, depositado no mesmo dia que o presente pedido, intitulado "*Audio Metadata Verification*," de Brett Graham Crockett, Registro do Procurador DOLI 50 e (3) e no desempenho da medição e correção de intensidade em um armazenamento de difusão ou cadeia de transmissão em que acesso ao áudio decodificado não é necessário e não é desejável. O referido N.S. 10/884.177 e os referidos pedidos com Registro de Procurador DOL150 são aqui incorporados através de referência em sua totalidade.

A economia de processamento proporcionada pelos aspectos da invenção também ajuda a tornar possível realizar medição de intensidade e a correção de metadados (por exemplo, mudando um parâmetro DIALNORM para o valor correto) em tempo real em um grande número de sinais de áudio comprimidos de dados de baixa taxa de bits. Frequentemente, muitos sinais de áudio codificados de baixa taxa de bits são multiplexados e transportados em fluxos de transporte de MPEG. A medição de intensidade de acordo com aspectos da presente invenção se torna medição de intensidade em tempo real em um grande número de sinais de áudio comprimidos muito mais possível, quando comparado com as exigências de decodificação com-

pleta dos sinais de áudio comprimidos para PCM a fim de realizar a medição de intensidade.

A figura 1 mostra uma disposição da técnica anterior para medição de intensidade de áudio codificado. Os dados de áudio codificados ou informação 101, tal como o áudio que foi codificado com baixa taxa de bits, são decodificados por um decodificador ou função de decodificação ("Decode" - Decodificar) 102 em, por exemplo, um sinal de áudio de PCM 103. Esse sinal é, então, aplicado a um medidor de intensidade ou método ou algoritmo de medição ("*Measure Loudness*" - Medir Intensidade) 104, que gera um valor de intensidade medido 105.

A figura 2 mostra um diagrama em blocos, estrutural ou funcional, da técnica anterior de um exemplo de um Decodificador 102. A estrutura ou funções que ele mostra são representativos de decodificadores Dolby Digital, Dolby Digital Plus e Dolby E. Quadros de dados de áudio codificados 101 são aplicados a um descompactador ou função de descompactação de dados ("*Frame Sync, Error Detection & Frame Deformatting*" - Sinc de Quadro, Detecção de Erro & Deformatação de Quadro) 202 que descompacta os dados aplicados em dados de expoente 203, dados de mantissa 204 e outra informação mista de alocação de bits 207. Os dados de expoente 203 são convertidos em um espectro de potência log 206 por um dispositivo ou função ("*Log Power Spectrum*" - Espectro de Potência Log) 205 e este espectro de potência log é usado por um alocador de bits ou função de alocação de bits ("*Bit Allocation*" Alocação de Bits) 208 para calcular sinal 209, que é o comprimento, em bits, de cada mantissa quantificada. As mantissas são, então, dequantificadas e combinadas com os expoentes por um dispositivo ou função ("*De-Quantize Mantissas*" - Dequantificar Mantissas) 210 e convertidas de volta para o domínio de tempo por um dispositivo ou função de banco de filtros inverso ("*Inverse Filterbank*" - Banco de Filtros Inverso) 212. Banco de filtros inverso corrente 212 também se sobrepõe e soma uma porção do resultado do Banco de Filtros Inverso corrente com o resultado do Banco de Filtros Inverso anterior (em tempo) para criar o sinal de áudio decodificado 103. Em implementações práticas de decodificar, recursos de

computação significativos são requeridos pelos dispositivos ou funções de Alocação de Bits, Dequantificação de Mantissas e Banco de Filtros Inverso. Mais detalhes do processo de decodificação podem ser encontrados em umas das referências citadas acima.

5 As figuras 3a e 3b mostram disposições da técnica anterior para medir, objetivamente, a intensidade de um sinal de áudio. Essas representam variações de Medir Intensidade 104 (figura 1). Embora as figuras 3a e 3b mostrem exemplos, respectivamente, de duas categorias gerais de técnicas de medição objetiva de intensidade, a escolha de uma técnica de medição objetiva particular não é crítica para a invenção e outras técnicas de medição objetiva de intensidade podem ser empregadas.

10 A figura 3a mostra um exemplo da disposição de medição de potência ponderada comumente usada na medição de intensidade. Um sinal de áudio 103 é passado através de uma função de filtro ou filtragem de ponderação ("Weighting Filter" - Filtro de Ponderação) 302, que é destinado a enfatizar frequências menos perceptivelmente sensíveis. A potência 305 do sinal filtrado 303 é calculada por um dispositivo ou função ("Potência") 304 e ponderada durante um período de tempo por um dispositivo ou função ("Média") 306 para criar um valor de intensidade 105. Existe um número de diferentes características de filtro de ponderação padrão e alguns exemplos comuns são mostrados na figura 4. Na prática, versões modificadas da disposição da figura 3a são usadas, freqüentemente, as modificações, por exemplo, impedindo períodos de tempo de silêncio de serem incluídos na média.

20 Técnicas baseadas em psicoacústica, freqüentemente, também são usadas para medir a intensidade. A figura 3b mostra uma disposição típica da técnica anterior dessa disposição baseada em psicoacústica. Um sinal de áudio 103 é filtrado por uma função de filtro ou filtragem de transmissão ("Transmission Filter" - Filtro de Transmissão) 312 que representa a resposta de magnitude de variação de frequência do ouvido externo e médio. O sinal filtrado 313, então, é separado por um banco de filtros ou função de banco de filtros auditivos ("Auditory Filterbank" - Banco de Filtros Auditivos) 314 em bandas de frequências que são equivalentes às bandas auditi-

vas críticas, ou mais estreitas do que as mesmas. Isso pode ser realizado através da realização de uma Transformação Rápida de Fourier (FFT) como implementado, por exemplo, por uma transformação discreta de frequências (DFT) e, então, agrupando as bandas espaçadas linearmente em bandas que se aproximam das bandas críticas do ouvido (com em uma escala de ERB ou Bark). Alternativamente, isso pode ser realizado por um único filtro de passagem de banda para cada banda de ERB ou Bark. Cada banda é, então, convertida por um dispositivo ou função "Excitação" 316 em um sinal de excitação 317, representando a quantidade de estímulos ou excitação experimentados pelo ouvido humano dentro da banda. A intensidade percebida ou a intensidade percebida para cada banda é, então, calculada da excitação por um dispositivo ou função ("*Specific Loudness*" - Intensidade Específica) 318 e a intensidade específica através de todas as bandas é somada por um somador ou função de soma ("*Sum*" - Soma) 320 para criar uma medição de intensidade 105 única. O processo de soma pode levar em consideração vários efeitos perceptuais, por exemplo, mascaramento de frequência. Em implementações práticas desses métodos perceptuais, recursos computacionais significativos são requeridos para o filtro de transmissão e banco de filtro auditivo.

20 A figura 5 mostra um diagrama em blocos de um aspecto da presente invenção. Um sinal de áudio digital codificado 101 é decodificado parcialmente por um dispositivo ou função decodificação parcial 502 e a intensidade é medida da informação parcialmente decodificada 503 por um dispositivo ou função medir intensidade 504. Dependendo de como a decodificação parcial é realizada, a medição de intensidade 505 resultante pode ser muito similar, mas não exatamente a mesma, à medição de intensidade 105, calculada do sinal de áudio completamente decodificado 103 (figura 1). No contexto de implementações de Dolby Digital, Dolby Digital Plus e Dolby E de aspectos da invenção, a decodificação parcial pode incluir a omissão dos dispositivos ou funções de Alocação de Bits, Dequantificação de Mantisas e Banco de Filtros Inverso de um decodificar, tal como o exemplo da figura 2.

As figuras 6a e 6b mostram dois exemplos de implementações da disposição geral da figura 5. Embora ambos possam empregar a mesma função ou dispositivo de decodificação parcial 502, cada um pode ter uma função ou dispositivo medir volume 504 diferente - que, no exemplo da figura 6a, é similar ao exemplo da figura 3a e que, no exemplo da figura 6b, é similar ao exemplo da figura 6b. Em ambos os exemplos, a Decodificação Parcial 502 extrai apenas os expoentes 203 do fluxo de áudio codificado e converte os expoentes em um espectro de potência 206. Essa extração pode ser realizada por um dispositivo ou função ("*Frame Sync, Error Detection & Frame De-Formatting*") 202 como no exemplo da figura 2 e essa conversão pode ser realizada por um dispositivo ou função ("*Low Power Spectrum*") 205 como no exemplo da figura 2. Não há exigência de de-quantificar as mantissas, realizar alocação de bits e realizar um banco de filtros inversos, como seria requerido para uma decodificação completa, conforme mostrado no exemplo de decodificação da figura 2.

O exemplo da figura 6a inclui uma Medição de Intensidade 504, que pode ser uma versão modificada do medidor de intensidade ou função de medição de intensidade da figura 3a. Neste exemplo, uma filtragem de ponderação modificada é aplicada no domínio de frequências pelo aumento ou diminuição dos valores de potência em cada banda por um filtro de ponderação ou função de filtragem ponderada ("*Modified Weighting Filter*") 601. Em contraste, o exemplo da figura 3a aplica filtragem de ponderação no domínio de tempo. Embora opere no domínio de frequências, o Filtro de Ponderação Modificada afeta o áudio da mesma maneira que o Filtro de Ponderação de domínio de tempo da figura 3a. O filtro 601 é "modificado" com relação ao filtro 302 da figura 3a no sentido de que opera em valores de amplitude log, em lugar de valores lineares e opera em uma escala de frequência não linear em lugar de linear. O espectro de potência ponderada de frequência 602 é, então, convertido em potência linear e somado através de frequência e a média determinada através do tempo por um dispositivo ou função - Converter, Somar, Tirar a Média 603 aplicando, por exemplo, a Equação 5, abaixo. A saída é um valor de intensidade objetivo 505.

O exemplo da figura 6b inclui uma Medição de Intensidade 504, que pode ser uma versão modificada do medidor de intensidade ou função de medição de intensidade da figura 3b. Neste exemplo, um filtro de transmissão ou função de filtragem modificada ("Filtro de Transmissão Modificado") 611 é aplicado diretamente no domínio de frequências pelo aumento ou diminuição dos valores de potência log em cada banda. Em contraste, o exemplo da figura 3b aplica filtragem de ponderação no domínio de tempo. Embora opere no domínio de frequências, o Filtro de Transmissão Modificado afeta o áudio da mesma maneira que o Filtro de Transmissão de domínio de tempo da figura 3b. Um banco de filtros ou função de banco de filtros auditivos modificados ("Banco de Filtros Auditivos Modificados") 613 aceita como entrada o espectro de potência log espaçado de banda de frequência linear e divide ou combina essas bandas linearmente espaçadas em uma saída de banco de filtros espaçados de banda crítica (por exemplo, bandas ERB ou Bark) 315. Banco de Filtros Auditivos Modificados 613 também converte o sinal de potência em um sinal linear para o dispositivo ou função de ("excitação") 316 seguinte. O Banco de Filtros Auditivos Modificados 613 é "modificado" com relação ao Banco de Filtros Auditivos 314 da figura 3b pelo fato de que opera em valores de amplitude log em lugar de valores lineares e converte esses valores de amplitude log em valores lineares. Alternativamente, o agrupamento de bandas em bandas ERB ou Bark pode ser realizado no Banco de Filtros Auditivos Modificados 613 em lugar do Filtro de Transmissão Modificado 611. O exemplo da figura 6b também inclui uma Intensidade Específica 318 para cada banda e uma Soma 320, como no exemplo a figura 3b.

Para as disposições mostradas nas figuras 6a e 6b, economia computacional significativa é obtida porque a decodificação não requer alocação de bits, dequantificação de mantissas e um banco de filtros inversos. Contudo, para as disposições da figura 6a e da figura 6b, a medição de intensidade objetiva resultante pode não ser exatamente a mesma que a medição calculada de áudio completamente decodificado. Isso é porque alguma da informação de áudio é descartada e, assim, a informação de áudio usada

para a medição é incompleta. Quando aspectos da presente invenção são aplicados ao Dolby Digital, Dolby Digital Plus ou Dolby E, a informação de mantissa é descartada e apenas os valores de expoentes grosseiramente quantificados são mantidos. Para Dolby Digital e Dolby Digital Plus os valores são quantificados em incrementos de 6 dB e para Dolby E eles são quantificados em incrementos de 3 dB. As etapas de quantificação menores em Dolby E resultam em valores de expoentes quantificados mais finos e, consequentemente, uma estimativa mais precisa do espectro de potência.

Os codificadores perceptuais, freqüentemente, são destinados a alterar o comprimento dos segmentos de tempo de sobreposição, também chamado o tamanho do bloco, em conjunto com certas características do sinal de áudio. Por exemplo, Dolby Digital usa dois tamanhos de blocos - um bloco mais longo de 512 amostras, predominantemente, para sinais de áudio estacionários e um bloco mais curto de 256 amostras para sinais de áudio mais transiente. O resultado é que o número de bandas de freqüência e o número correspondente de valores de espectro de potência log 206 varia bloco a bloco. Quando o tamanho do bloco é 512 amostras, há 256 bandas e, quando o tamanho de bloco é 256 amostras, há 128 bandas.

Há muitas maneiras pelas quais os métodos propostos nas figuras 6a e 6b podem manipular blocos de tamanhos variáveis e cada maneira leva a uma medição de intensidade resultante similar. Por exemplo, o Espectro de Potência Log 205 pode ser modificado para sair sempre um número constante de bandas em uma taxa de blocos constante por meio da combinação ou da média de múltiplos blocos menores em blocos maiores e dispersando a potência do número menor de bandas através do número maior de bandas. Alternativamente, a Medição de Intensidade pode aceitar tamanhos de blocos variáveis e ajustar, consequentemente, seus processos de filtragem, excitação, intensidade específica, determinação de média e soma, por exemplo, pelo ajuste das constantes de tempo.

Exemplo de Medição de Potência Ponderada

Como um exemplo de aspectos da presente invenção, uma versão altamente econômica de um método de medição de intensidade de po-

tência ponderada pode usar fluxos de bits de Dolby Digital e a medição de intensidade de potência ponderada LeqA. Neste exemplo altamente econômico, apenas os expoentes quantificados contidos em um fluxo de bits Dolby Digital são usados como uma estimativa do espectro de sinal de áudio para realizar a medição de intensidade. Isso evita as exigências computacionais adicionais de realização de alocação de bits para recrear a informação de mantissa, o que, caso contrário, proporcionaria apenas uma estimativa ligeiramente mais precisa do espectro do sinal.

Conforme representado nos exemplos das figuras 5 e 6a, o fluxo de bits é parcialmente decodificado para recriar e extrair o espectro de potência log, calculado dos dados de expoente quantificados contidos no fluxo de bits. O Dolby Digital realiza codificação de áudio de baixa taxa de bits através de formação de janelas 512 consecutivas, 50% de amostras de áudio de PCM sobrepostas e realizando uma transformação de MDCT, resultando em 256 coeficientes de MDCT que são usados para criar o fluxo de áudio codificado de baixa taxa de bits. A decodificação parcial realizada nas figuras 5 e 6a descompacta os dados de expoentes $E(k)$ e converte os dados descompactados em 256 valores de espectro de potência log quantificados, $P(k)$, que forma uma representação espectral bruta do sinal de áudio. Os valores de espectro de potência log, $P(k)$, estão em unidades de dB. A conversão é como segue:

$$P(k) = -E(k) \cdot 20 \cdot \log_{10}(2) \quad 0 \leq k < N \quad (1)$$

onde $N = 256$, o número de coeficientes de transformação para cada bloco em um fluxo de bits Dolby Digital. Para usar o espectro de potência log na computação da medição de intensidade de potência ponderada, o espectro de potência log é ponderado usando uma curva de intensidade apropriada, tal como uma das curvas de ponderação A-, B- ou C-, mostradas na figura 4. Nesse caso, a medida de potência de LeqA está sendo computada e, portanto, a curva de ponderação A- é apropriada. Os valores de espectro de potência log $P(k)$ são ponderados pela adição dos mesmos aos valores de frequências de ponderação de A-, $A_w(k)$, também em unidades de dB como

$$P_{pw}(k) = P(k) + A_{pw}(k) \quad 0 \leq k < N \quad (2)$$

Os valores de frequências discretas de ponderação de A-, $A_w(k)$, são criados pela computação dos valores de ganho de ponderação de A- para as frequências distintas, f_{distinta} , onde

$$f_{\text{discreta}} = \frac{F}{2} + F \cdot k \quad 0 \leq k < N \quad (3)$$

Onde

$$F = \frac{F_s}{2 \cdot N} \quad 0 \leq k < N \quad (4)$$

- 5 e onde a frequência de amostragem F_s é, tipicamente, igual a 48 kHz para Dolby Digital. Cada conjunto de valores de espectro de potência log ponderada, $P_w(k)$ é, então, convertido de dB para potência linear e somado para criar estimativa de potência ponderada de A- P_{POW} das 512 amostras de áudio de PCM como

$$P_{POW} = \sum_{k=0}^{N-1} 10^{(P_w(k)/10)} \quad (5)$$

- 10 Conforme mencionado previamente, cada fluxo de bits de Dolby Digital contém transformações consecutivas criadas por formação de janelas de 512 amostras de PCM com 50% de sobreposição e realização da transformação de MDCT. Portanto, uma aproximação da potência ponderada de A- total, P_{TOT} , da baixa taxa de bits de áudio codificados em um fluxo de bits
- 15 Dolby Digital pode ser computada por meio da média dos valores de potência através de todas as transformações no fluxo de bits Dolby Digital como segue:

$$P_{TOT} = \frac{1}{M} \sum_{m=0}^{M-1} P_{POW}(m) \quad (6)$$

- onde M iguala o número total de transformações contidas no fluxo de bits Dolby Digital. A potência média é, então, convertida em unidades de dB como segue:
- 20

$$L_A = 10 \cdot \log_{10}(P_{TOT}) - C \quad (7)$$

onde C é um deslocamento constante devido às mudanças de nível realizadas no processo de transformação durante codificação do fluxo de bits Dolby Digital.

Exemplo de Medição Psicoacústica

Como outro exemplo de aspectos da presente invenção, uma versão altamente econômica de um método de medição de intensidade de potência ponderada pode usar fluxos de bits Dolby Digital e uma medição de intensidade psicoacústica. Nesse exemplo altamente econômico, como no anterior, apenas os expoentes quantificados contidos em um fluxo de bits Dolby Digital são usados como uma estimativa do espectro do sinal de áudio para realizar a medição de intensidade. Como no outro exemplo, isso evita as exigências computacionais adicionais de realização de alocação de bits para recriar a informação de mantissa, que, de outro modo, apenas proporcionaria uma estimativa ligeiramente mais precisa do espectro do sinal.

O Pedido de Patente Internacional Nº PCT/US2004/016964, depositado em 27 de maio de 2004, Seefeldt e *outros*, publicado como WO 2004/111994 A2, 23 de dezembro de 2004, pedido que designa os Estados Unidos, descreve, entre outras coisas, uma medida objetiva de intensidade percebida com base em um modelo psicoacústico. O referido pedido é aqui incorporado através de referência em sua totalidade. Os valores de espectro de potência log, $P(k)$, derivados da decodificação parcial de um fluxo de bits Dolby Digital, podem servir como entradas para uma técnica, tal como no referido pedido internacional, bem como outras medidas psicoacústicas similares, em lugar do áudio de PCM original. Essa disposição é mostrada no exemplo da figura 6b. Tomando emprestada a terminologia e a notação do referido pedido de PCT, um sinal de excitação $E(b)$ aproximando-se da distribuição de energia ao longo da membrana basilar do ouvido interno na banda crítica b pode ser aproximada dos valores de espectro de potência log como segue:

$$E(b) = \sum_k |T(k)|^2 |H_b(k)|^2 10^{P(k)/10} \quad (8)$$

onde $T(k)$ representa a resposta de frequência do filtro de transmissão e $H_b(k)$ representa a resposta de frequência da membrana basilar em uma localização correspondente à banda crítica b , ambas as respostas sendo amostradas na frequência correspondente ao bin de transformação k . Em seguida, as excitações correspondentes a todas as transformações no fluxo de bits Dolby Digital têm suas médias determinadas para produzir uma excita-

ção total:

$$\bar{E}(b) = \frac{1}{M} \sum_m E(b, m) \quad (9)$$

Usando contornos de intensidade iguais, a excitação total em cada banda é transformada em um nível de excitação que gera a mesma intensidade em 1 kHz. Intensidade específica, uma medição de intensidade perceptual distribuída através da frequência, é, então, computado da excitação transformada, $\bar{E}_{1kHz}(b)$, através de uma linearidade não compressiva:

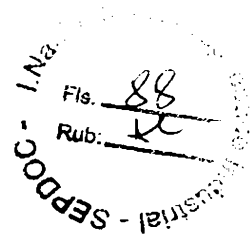
$$N(b) = G \left(\left(\frac{\bar{E}_{1kHz}(b)}{TO_{1kHz}} \right)^\alpha - 1 \right) \quad (10)$$

onde TO_{1kHz} é o limite em silêncio em 1kHz e as constantes G e α são escolhidas para corresponder aos dados gerados dos experimentos psicoacústicos descrevendo o crescimento da intensidade. Finalmente, a intensidade total, L , representada em unidades de sone, é computada pela soma da intensidade específica através de bandas:

$$L \approx \sum_b N(b) \quad (11)$$

Para fins de ajuste do sinal de áudio, pode-se desejar computar um ganho correspondente, $G_{\text{correspondência}}$, que, quando multiplicado com o sinal de áudio torna a intensidade do áudio ajustado igual a uma intensidade de referência, L_{REF} , como medida pela técnica psicoacústica descrita. Como a medição psicoacústica envolve uma não linearidade na computação de intensidade específica, uma solução de forma fechada para $G_{\text{correspondência}}$ não existe. Na verdade, uma técnica iterativa descrita no referido pedido de PCT pode ser empregada em que o quadrado do ganho correspondente é ajustado e multiplicado com a excitação total, $\bar{E}(b)$, até que a intensidade total correspondente, L , esteja dentro de uma diferença de limite com relação à intensidade de referência, L_{REF} . A intensidade do áudio pode, então, ser expressa em dB com relação à referência como:

$$L_{dB} = 20 \log_{10} \left(\frac{1}{G_{\text{correspondência}}} \right) \quad (12)$$



Outros Codecs de Áudio Perceptíveis

Aspectos da presente invenção não estão limitados aos sistemas de codificação Dolby Digital, Dolby Digital Plus e Dolby E. Sinais de áudio codificados, usando certos outros sistemas de codificação, em que uma aproximação do espectro de potência do áudio é proporcionada, por exemplo, fatores de escala, envelopes espectrais e coeficientes preditivos lineares, que podem ser recuperados de um fluxo de bits codificado, sem decodificar completamente o fluxo de bits para produzir áudio, também podem se beneficiar de aspectos da presente invenção.

10 Erro no Cálculo de Potência de Expoentes Dolby Digital

Os expoentes Dolby Digital $E(k)$ representam uma quantificação bruta do logaritmo dos coeficientes de espectro de MDCT. Há um número de fontes de erro, quando usando esses valores como um espectro de potência bruto.

15 Primeiro, no Dolby Digital, o próprio processo de quantificação resulta em erro médio de, aproximadamente, 2,7 dB, quando comparando os valores do espectro de potência gerado dos expoentes (veja equação 1, acima) e os valores de potência, calculados diretamente dos coeficientes de MDCT. Esse erro médio, que foi determinado, experimentalmente, pode ser incorporado no deslocamento constante C na Equação 7, acima.

20 Em segundo lugar, sob certas condições de sinal, tais como transientes, valores de expoentes são agrupados através de frequência (referido como modos "D25" e "D45", no documento A/52A, citado acima). Esse agrupamento através de frequência faz com que o erro médio de expoente seja menos predizível e, assim, mais difícil de levar em conta pela incorporação na constante C da Equação 7. Na prática, o erro devido a esse agrupamento pode ser ignorado por duas razões: (1) o agrupamento é usado raramente e (2) a natureza dos sinais para os quais o agrupamento é usado resulta em um erro médio medido, que é similar ao caso de média não determinada.

30

Implementação

A invenção pode ser implementada em *hardware* ou *software*, ou

uma combinação de ambos (por exemplo, arranjos lógicos programáveis). A menos que de outro modo especificado, os algoritmos ou processos incluídos como parte da invenção não estão, inerentemente, relacionados com qualquer computador particular ou outro aparelho. Em particular, várias máquinas para fins gerais podem ser usadas como programas escritos de acordo com os ensinamentos aqui, ou pode ser mais conveniente construir aparelho mais especializado (por exemplo, circuitos integrados) para realizar as etapas do método requeridas. Desse modo, a invenção pode ser implementada em um ou mais programas de computador, executando em um ou mais sistemas de computador programáveis, cada um compreendendo pelo menos um processador, pelo menos um sistema de armazenamento de dados (incluindo memória não volátil e/ ou elementos de armazenamento), pelo menos um dispositivo ou porta de entrada e pelo menos um dispositivo ou porta de saída. O código de programa é aplicado aos dados de entrada para realizar as funções aqui descritas e gerar informação de saída. A informação de saída é aplicada a um ou mais dispositivos de saída, de modo conhecido.

Cada um desses programas pode ser implementado em qualquer linguagem de computador desejada (incluindo máquina, montagem, ou procedimento de nível elevado, lógica ou linguagens de programação orientadas em objeto) para se comunicar com um sistema de computador. Em qualquer caso, a linguagem pode ser uma linguagem compilada ou interpretada.

Será apreciado que algumas etapas ou funções mostradas nas figuras exemplificativas realizam múltiplas subetapas e também podem ser mostradas como múltiplas etapas ou funções, em lugar de uma etapa ou função. Também será apreciado que vários dispositivos, funções, etapas e processos mostrados e descritos aqui em vários exemplos podem ser mostrados combinados ou separados de outras maneiras que não conforme mostrado nas várias figuras. Por exemplo, quando implementadas por seqüências de instrução de *software* de computador, várias funções e etapas das figuras exemplificativas podem ser implementadas por seqüências de instrução de *software* multiencadeadas, executando em hardware de pro-

cessamento de sinal digital adequado, em cujo caso os vários dispositivos e funções nos exemplos mostrados nas figuras pode corresponder às porções das instruções de *software*.

5 Cada programa desse computador, de preferência, é armazenado ou baixado para um meio ou dispositivo de armazenamento (por exemplo, memória de estado sólido ou média, ou meios magnéticos ou óticos) legíveis por um computador programável para fins gerais ou especiais para
10 configuração e operação do computador quando os meios ou dispositivo de armazenamento é lido pelo sistema de computador para realizar os procedimentos aqui descritos. O sistema da invenção também pode ser considerado para ser implementado como um meio de armazenamento legível em computador, configurado com um programa de computador, onde o meio de armazenamento assim configurado faz com que um sistema de computador opere em uma maneira específica e predeterminada para realizar as funções
15 aqui descritas.

Um número de modalidades da invenção foi descrito. Não obstante, será compreendido que várias modificações podem ser feitas, sem afastamento do espírito e do escopo da invenção. Por exemplo, algumas das etapas aqui descritas podem ser independentes de ordem e, assim, pode ser
20 realizado em uma ordem diferente daquela descrita.

REIVINDICAÇÕES

1. Método para medir a intensidade de áudio codificado em um fluxo de dados, **caracterizado** por compreender as seguintes etapas:

5 aproximar o espectro de potência do áudio sem decodificar completamente o áudio; e

aproximar a intensidade do áudio a partir da aproximação do espectro de potência.

2. Método, de acordo com a reivindicação 1, **caracterizado** pelo fato de que os referidos dados incluem representações brutas do áudio e representações mais finas associadas do áudio, e

10 em que a referida aproximação do espectro de potência do áudio é derivada das representações brutas do áudio.

3. Método, de acordo com a reivindicação 2, **caracterizado** pelo fato de que o áudio é sub-banda codificada, usando sub-bandas de frequência, cada uma das sub-bandas de frequência tendo um fator de escala e dados de amostra associados com os mesmos, e

15 em que as representações brutas do áudio compreendem fatores de escala e as representações mais finas associadas do áudio compreendem dados de amostra associados com cada fator de escala.

20 4. Método, de acordo com a reivindicação 3, **caracterizado** pelo fato de que o fator de escala e os dados de amostra de cada sub-banda representam coeficientes espectrais na sub-banda pela notação exponencial em que o fator de escala compreende um expoente e os dados de amostra associados compreendem mantissas.

25 5. Método, de acordo com qualquer uma das reivindicações de 1 a 4, **caracterizado** pelo fato de que o áudio é AC-3 codificado.

6. Método, de acordo com a reivindicação 2, **caracterizado** pelo fato de que o áudio codificado é codificado preditivo linear onde as representações brutas do áudio compreendem coeficientes preditivos lineares, e

30 as representações mais finas do áudio compreendem informação de excitação associada com os coeficientes preditivos lineares.

7. Método, de acordo com a reivindicação 2, **caracterizado** pelo

fato de que as representações brutas do áudio compreendem um envelope espectral, e

as representações mais finas do áudio compreendem componentes espectrais associados com o envelope espectral.

5 8. Método, de acordo com qualquer uma das reivindicações de 1 a 7, **caracterizado** pelo fato de que aproximar o áudio inclui

aplicar uma medição de intensidade de potência ponderada.

9. Método, de acordo com a reivindicação 8, **caracterizado** pelo fato de que a medição de intensidade de potência ponderada emprega um
10 filtro que tira ênfase de frequências menos perceptíveis e determina a média da potência do áudio filtrado ao longo do tempo.

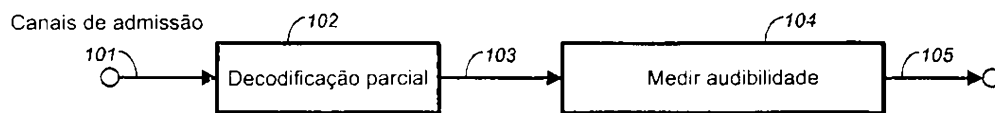
10. Método, de acordo com qualquer uma das reivindicações de 1 a 7, **caracterizado** pelo fato de que aproximar a intensidade do áudio em resposta à aproximação do espectro de potência do áudio inclui

15 aplicar uma medição de intensidade psicoacústica.

11. Método, de acordo com a reivindicação 10, **caracterizado** pelo fato de que a medição de intensidade psicoacústica emprega um modelo do ouvido humano para determinar intensidade específica em cada uma de uma pluralidade de bandas de frequência similares às bandas críticas do
20 ouvido humano.

12. Método, de acordo com a reivindicação 10 como dependente da reivindicação 3 ou 4 **caracterizado** pelo fato de que as sub-bandas são similares às bandas críticas do ouvido humano e

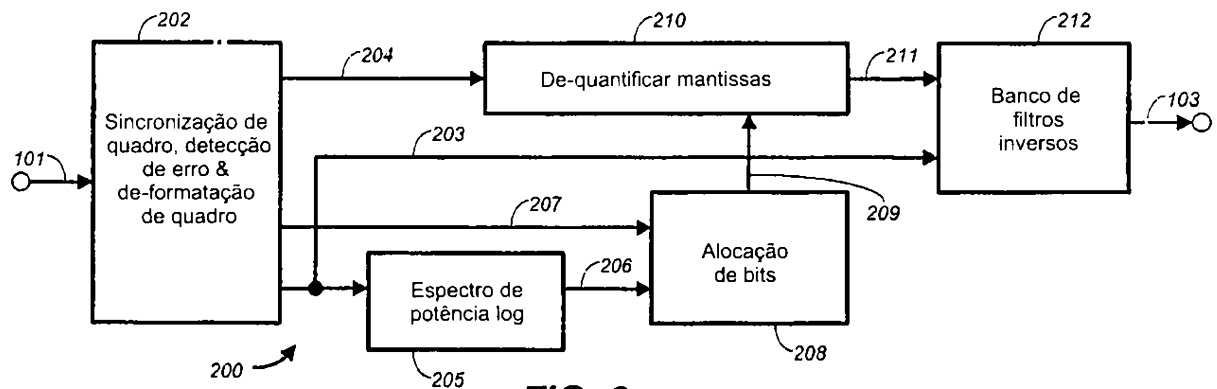
a medição de intensidade psicoacústica emprega um modelo do
25 ouvido humano para determinar intensidade específica em cada uma das sub-bandas.



100 ↗

FIG. 1

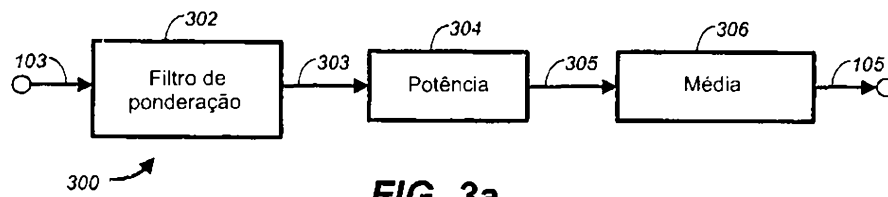
Técnica Anterior



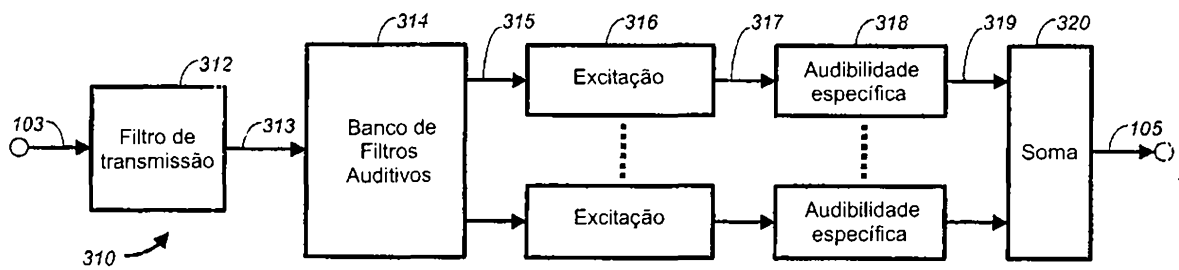
200 ↗

FIG. 2

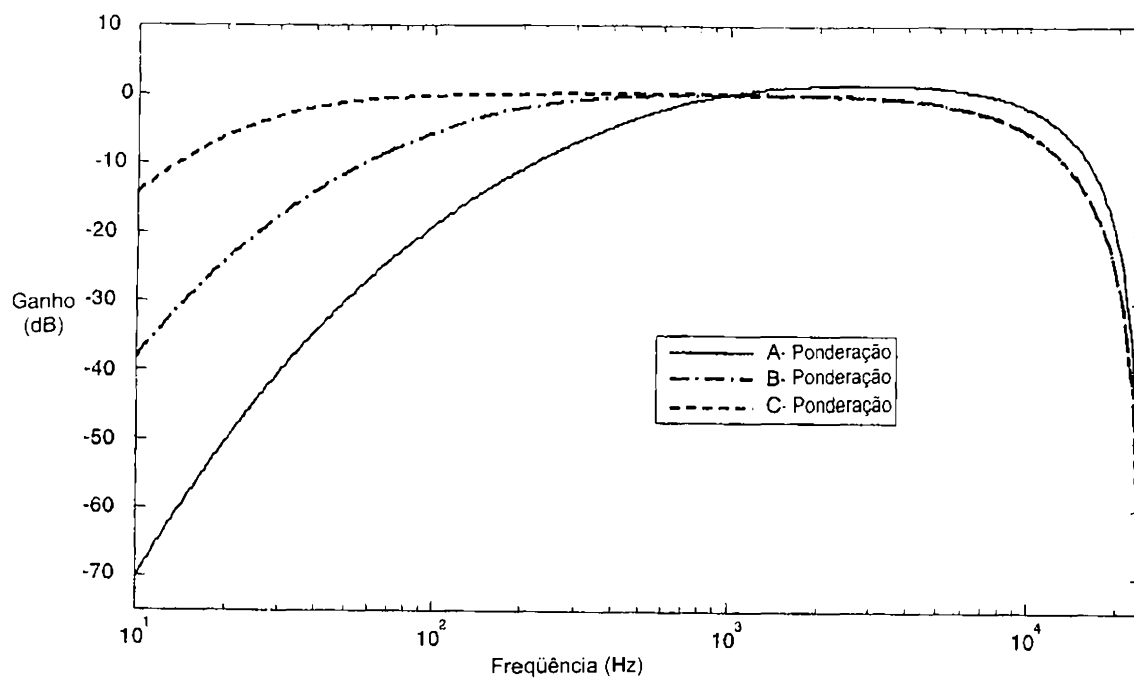
Técnica Anterior

**FIG. 3a**

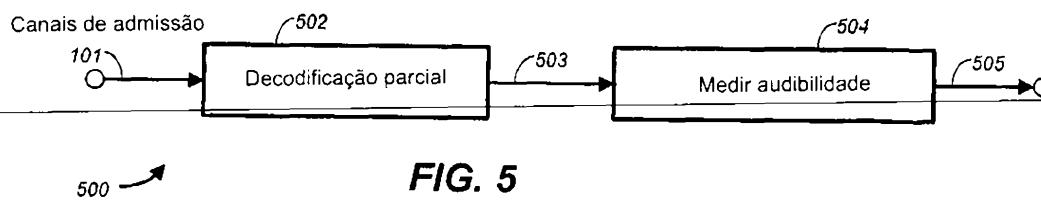
Técnica Anterior

**FIG. 3b**

Técnica Anterior

**FIG. 4**

Técnica Anterior

**FIG. 5**

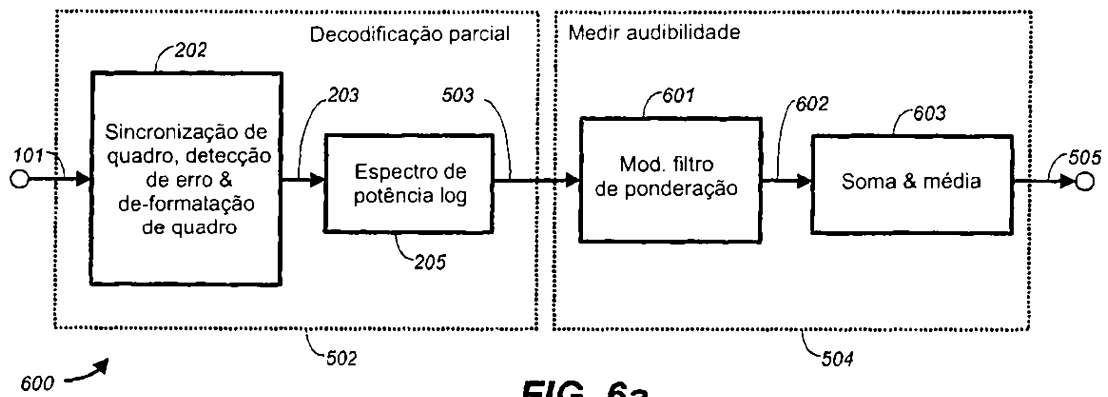


FIG. 6a

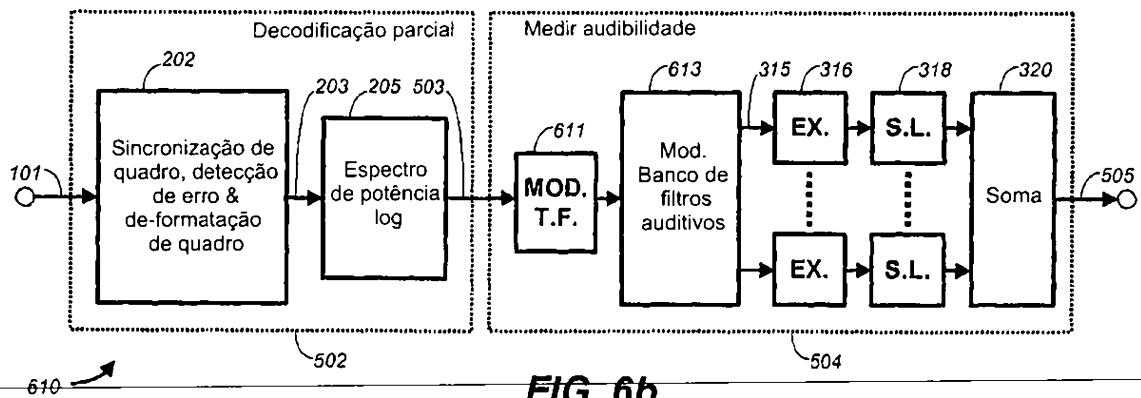


FIG. 6b