



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

(11) PI 0709877-4 B1



(22) Data do Depósito: 30/03/2007

(45) Data de Concessão: 31/12/2019

(54) Título: MÉTODO E APARELHO PARA CONTROLAR UMA CARACTERÍSTICA DE INTENSIDADE ACÚSTICA PARTICULAR DE UM SINAL DE ÁUDIO

(51) Int.Cl.: G10L 25/69; G10L 21/0364; H03G 5/16; H03G 9/00; G10L 21/0264; (...).

(52) CPC: G10L 25/69; G10L 21/0364; H03G 5/165; H03G 9/005; G10L 21/0264; (...).

(30) Prioridade Unionista: 04/04/2006 US 60/789,539.

(73) Titular(es): DOLBY LABORATORIES LICENSING CORPORATION.

(72) Inventor(es): ALAN JEFFREY SEEFELDT.

(86) Pedido PCT: PCT US2007007946 de 30/03/2007

(87) Publicação PCT: WO 2007/120453 de 25/10/2007

(85) Data do Início da Fase Nacional: 03/10/2008

(57) Resumo: CÁLCULO E AJUSTE DA INTENSIDADE ACÚSTICA PERCEBIDA E/OU DO EQUILÍBRIO ESPECTRAL PERCEBIDO DE UM SINAL DE ÁUDIO Um processamento de sinal de áudio relativo à medição e controle da intensidade acústica de som percebida e/ou do equilíbrio espectral percebido de um sinal de áudio é útil, por exemplo, em um ou mais dentre: o controle do volume compensado em intensidade acústica, o controle do ganho automático, o controle da faixa dinâmica (incluindo, por exemplo, limitadores, compressores, expansores, etc.), a equalização dinâmica, e a compensação para a interferência de ruído de fundo em um ambiente de reprodução de áudio. Em várias modalidades, parâmetros de modificação são obtidos de modo a modificar o sinal de áudio a fim de reduzir a diferença entre a sua intensidade acústica específica e a intensidade acústica alvo específica.

Relatório Descritivo da Patente de Invenção para "**MÉTODO E APARELHO PARA CONTROLAR UMA CARACTERÍSTICA DE INTENSIDADE ACÚSTICA PARTICULAR DE UM SINAL DE ÁUDIO**".

CAMPO DA TÉCNICA

[001] A presente invenção refere-se a um processamento de sinal de áudio. Mais particularmente, a presente invenção se refere à medição e ao controle da intensidade acústica do som percebida e/ou equilíbrio espectral percebido de um sinal de áudio. A presente invenção é útil, por exemplo, em um ou mais dentre: o controle do volume compensado em intensidade acústica, o controle do ganho automático, o controle da faixa dinâmica (incluindo, por exemplo, limitadores, compressores, expansores, etc.), a equalização dinâmica, e a compensação para a interferência de ruído de fundo em um ambiente de reprodução de áudio. A presente invenção inclui não apenas métodos, mas também os programas e aparelhos de computador correspondentes.

ANTECEDENTES DA INVENÇÃO

[002] Têm-se feito muitas tentativas no sentido de desenvolver um método objetivo satisfatório de se medir uma intensidade acústica. Em 1933, Fletcher e Munson determinaram que a audição humana é menos sensível a baixas e altas frequências que a frequências médias (ou da voz). Também descobriram-se que a mudança relativa de sensibilidade diminuía conforme o nível do som aumentava. Um medidor de intensidade acústica da técnica anterior consiste em um microfone, um amplificador, um medidor, e de uma combinação de filtros desenhados para imitar grosseiramente a resposta de frequência da audição nos níveis baixo, médio e alto de som.

[003] Mesmo que tais dispositivos provessem uma medição da intensidade acústica de um único nível constante, as medições tonais isoladas de sons mais complexos não corresponderiam muito bem às

impressões subjetivas da intensidade acústica. Os medidores de níveis de som deste tipo foram padronizados, mas só são usados para tarefas específicas, tais como para o monitoramento e controle do barulho industrial.

[004] No início dos anos 50, Zwicker e Stevens, entre outros, estenderam o trabalho de Fletcher e Munson no desenvolvimento de um modelo mais realista do processo de percepção da intensidade acústica. Em 1956, Stevens publicou um método para o "Cálculo da Intensidade Acústica de Ruídos Complexos" no Periódico da Sociedade Acústica da América, e, em 1958, Zwicker publicou o seu artigo sobre Acústica intitulado "Base Psicológica e Metodológica da Intensidade Acústica". Em 1959, Zwicker publicou um procedimento gráfico para o cálculo da intensidade acústica, assim como vários artigos similares logo em seguida. Os métodos de Stevens e de Zwicker foram padronizados como ISO 532, partes A e B (respectivamente). Ambos os métodos envolvem etapas similares.

[005] Primeiramente, a distribuição variante no tempo da energia ao longo da membrana basilar do ouvido interno, referida como excitação, é simulada ao passar um áudio por um banco de filtros auditivos de passa banda com frequências centrais espaçadas uniformemente em uma escala de taxa de banda crítica. Cada filtro auditivo é concebido de modo a simular a resposta de frequência em um local em particular ao longo da membrana basilar do ouvido interno, com a frequência central do filtro correspondente a este local. Uma largura de banda crítica é definida como a largura de banda de tal filtro. Medida em unidades de Hertz, a largura de banda crítica destes filtros auditivos aumenta com uma maior frequência central. É, portanto, útil se definir uma escala de frequência deformada de tal modo que a largura de banda crítica para todos os filtros auditivos medidos nesta escala deformada seja constante. Tal escala deformada é referida como uma

escala de taxa de banda crítica e é muito útil para o entendimento e simulação de uma ampla faixa de fenômenos psicoacústicos. Vide, por exemplo, *Psychoacoustics - Facts and Models* de E. Zwicker e H. Fastl, Springer-Verlag, Berlim, 1990. Os métodos de Stevens e Zwicker utilizam uma escala de taxa de banda crítica referida como escala de Bark, na qual a largura de banda crítica é constante abaixo de 500 Hz e aumenta acima de 500 Hz. Mais recentemente, Moore e Glasberg definiram uma escala de taxa de banda crítica, que nomearam escala de Largura de Banda Retangular Equivalente (ERB) (B.C.J. Moore, B. Glasberg, T. Baer, "A Model for the Prediction of Thresholds, Loudness, and Partial Loudness", *Journal of the Audio Engineering Society*, Vol. 45, N. 4. abril de 1997, pp. 224-240). Através de experimentos psicoacústicos que utilizam mascarados de ruído chanfrados, Moore e Glasberg demonstraram que a largura de banda crítica continua a diminuir abaixo de 500 Hz em contraste à escala de Bark na qual a largura de banda crítica permanece constante.

[006] Após a computação da excitação tem-se uma função compressiva não linear que gera uma quantidade referida como "intensidade acústica específica". A intensidade acústica específica é uma medida da intensidade acústica perceptiva como uma função de frequência e do tempo e pode ser medida em unidades de intensidade acústica perceptiva por frequência unitária ao longo de uma escala de taxa de banda crítica, tal como a escala de Bark ou de banda ERB acima apresentada. Em termos ideais, a intensidade acústica específica representa uma distribuição contínua de intensidade acústica como uma função de frequência e de tempo, e a "intensidade acústica total" variante no tempo é computada ao integrar esta distribuição através da frequência. Na prática, uma representação precisa de uma intensidade acústica específica é obtida por meio da amostragem desta distribuição de maneira uniforme ao longo de uma escala de taxa de banda

crítica através do uso, por exemplo, dos filtros auditivos acima mencionados. Neste caso, a intensidade acústica total pode ser computada ao simplesmente somar a intensidade acústica específica de cada filtro. A fim de reduzir complexidade, algumas aplicações podem computar uma aproximação mais grosseira da intensidade acústica específica graças às pequenas imprecisões na estimativa e na modificação da intensidade acústica percebida. Tais aproximações serão apresentadas em mais detalhes mais adiante.

[007] A intensidade acústica pode ser medida em unidades de phon. A intensidade acústica de um dado som em phon é o nível de pressão de som (SPL) de um tom de 1 kHz tendo uma intensidade acústica subjetiva igual ao daquele som. Convencionalmente, a referência 0 dB para o nível SPL é uma pressão de raiz quadrada mínima de 2×10^{-5} Pascal, e isto é, também, portanto, a referência 0 phon. Usando esta definição em comparação à intensidade acústica dos tons em frequências diferentes de 1 kHz com a intensidade acústica a 1 kHz, um contorno de intensidade acústica igual poderá ser determinado para um dado nível de phon. A Figura 11 mostra contornos de intensidade acústica iguais para frequências entre 20 Hz e 12,5 kHz, e para níveis de phon entre 4,2 phon (considerado como o limite da audição) e 120 phon (ISO226: 1087 (E), "Acoustics - Normal equal loudness level contours"). A medição de phon leva em consideração a variante sensibilidade do ouvido humano à frequência, mas os resultados não permitem a avaliação das intensidades acústicas subjetivas relativas dos sons em níveis variantes, uma vez que não se faz uma tentativa no sentido de corrigir a não linearidade do crescimento da intensidade acústica com nível SPL, ou seja, para o fato de que o espaçamento dos contornos varia.

[008] A intensidade acústica pode também ser medida em unidades de "sone". Existe um mapeamento de um para um entre as unida-

des de phon e as unidades de sone, conforme indicado na Figura 11. Um som é definido como a intensidade acústica de uma onda seno pura de 1 kHz 40 dB (SPL) e equivale a 40 phon. As unidades de sone são tais que um aumento dobrado em sone corresponderá ao dobro da intensidade acústica percebida. Por exemplo, 4 sones são percebidos como duas vezes tão alto quanto 2 sons. Sendo assim, os níveis de intensidade acústica expressos em sone são mais informativos. Dada a definição da intensidade acústica como uma medida da intensidade acústica perceptiva como uma função de frequência e tempo, uma intensidade acústica específica poderá ser medida em unidades de sone por frequência unitária. Deste modo, ao se usar a escala de Bark, a intensidade acústica específica tem unidades de sone por Bark e, da mesma forma, ao se usar a escala ERB, as unidades serão em sone por ERB.

[009] Conforme acima mencionado, a sensibilidade do ouvido humano varia de acordo com a frequência e o nível, um fato bem documentado na literatura psicoacústica. Um dos resultados é que o espectro ou timbre percebido de um dado som varia de acordo com o nível acústico no qual o som é ouvido. Por exemplo, para um som contendo frequências baixas, médias e altas, as proporções relativas percebidas de tais componentes de frequência mudam de acordo com a intensidade acústica geral do som; quando o som é calmo, os componentes de frequências baixa e alta soam mais calmos com relação às frequências médias do que, quando o som é alto. Este fenômeno é bem conhecido e é suavizado em um equipamento de reprodução de som por meio dos assim chamados controles de intensidade acústica. Um controle de intensidade acústica é um controle de volume que aplica um reforço (boost) de frequência baixa e às vezes também alta quando o volume é desligado. Sendo assim, a sensibilidade menor do ouvido em frequências extremas é compensada por um reforço artifici-

al destas frequências. Estes controles são completamente passivos; o grau de compensação aplicado é uma função da definição do controle de volume ou algum outro controle operado pelo usuário, não como uma função do conteúdo dos sinais de áudio.

[0010] Na prática, as mudanças no equilíbrio espectral relativo percebido entre frequências baixa, média e alta dependem do sinal, em particular de seu espectro real e de se quer ser alto ou suave. Considere a gravação de uma orquestra sinfônica. Reproduzido no mesmo nível que uma pessoa da plateia ouviria em um concerto, o equilíbrio através do espectro poderá ser correto se a orquestra estiver tocando alto ou baixo. Se a música é reproduzida 10 dB mais baixo, por exemplo, o equilíbrio percebido através do espectro muda de uma maneira nas passagens altas e muda de uma outra maneira nas passagens baixas. Um controle de intensidade acústica passiva convencional não aplica diferentes compensações como uma função da música.

[0011] No Pedido de Patente Internacional Nº PCT/US2004/016964, depositado em 27 de maio de 2004, publicado em 23 de dezembro de 2004 como o documento WO 2004/111994 A2, Seefeldt e outros apresentam, entre outras coisas, um sistema para a medição e ajuste da intensidade acústica percebida de um sinal de áudio. O dito pedido PCT, que designa os Estados Unidos da América, é incorporado ao presente documento à guisa de referência em sua totalidade. No dito pedido, um modelo psicoacústico calcula a intensidade acústica de um sinal de áudio em unidades perceptivas. Além disso, o pedido introduz técnicas para a computação de um ganho multiplicativo de banda larga que, quando aplicado ao áudio, resulta em uma intensidade acústica do áudio modificado no ganho substancialmente igual como uma intensidade acústica de referência. A aplicação de tal ganho de banda larga, no entanto, muda o equilíbrio es-

pectral do áudio.

DESCRIÇÃO DA INVENÇÃO

[0012] Em um aspecto, a presente invenção provê a obtenção de informações utilizáveis para o controle da intensidade acústica específica de um sinal de áudio por meio da modificação do sinal de áudio a fim de reduzir a diferença entre a sua intensidade acústica específica e uma intensidade acústica alvo específica. A intensidade acústica específica é uma medida da intensidade acústica perceptiva como uma função da frequência e do tempo. Em implementações práticas, a intensidade acústica específica do sinal de áudio modificado pode ser feita de modo a se aproximar da intensidade acústica alvo específica. A aproximação pode ser afetada não apenas por considerações simples de processamento de sinal, mas também pela suavização no tempo e/ou na frequência que pode ser empregada na modificação, conforme descrito abaixo.

[0013] Uma vez que a intensidade acústica específica é uma medida da intensidade acústica perceptiva de um sinal de áudio como uma função da frequência e do tempo, a fim de reduzir a diferença entre a intensidade acústica específica do sinal de áudio e a intensidade acústica alvo específica, a modificação pode modificar o sinal de áudio como uma função da frequência. Embora em alguns casos a intensidade acústica alvo específica possa ser invariante no tempo e o próprio sinal de áudio possa ser um sinal invariante no tempo de estado estacionário, tipicamente, a modificação pode também modificar o sinal de áudio como uma função do tempo.

[0014] Aspectos da presente invenção podem também ser empregados no sentido de compensar a interferência do ruído de fundo em um ambiente de reprodução de áudio. Quando um áudio é ouvido na presença de um ruído de fundo, o ruído pode parcial ou completamente mascarar o áudio de maneira dependente tanto no nível ou espectro

do áudio, como também no nível ou espectro do ruído. O resultado é uma alteração do espectro percebido do áudio. De acordo com estudos psicoacústicos (vide, por exemplo, Moore, Glasberg, e Baer, "A Modelo for the Prediction of Thresholds, Loudness, and Partial Loudness", *J. Audio Eng. Soc.*, Vol. 45, N. 4, abril de 1997), pode-se definir a intensidade acústica parcial específica" do áudio como a intensidade acústica perceptiva do áudio na presença de um sinal sonoro de interferência secundário, como o ruído.

[0015] Sendo assim, em um outro aspecto, a presente invenção provê a obtenção de informações utilizáveis para o controle da intensidade acústica parcial específica de um sinal de áudio por meio da modificação de um sinal de áudio a fim de reduzir a diferença entre a sua intensidade acústica parcial específica e uma intensidade acústica alvo específica. Isto suaviza os efeitos do ruído de uma maneira perceptivamente precisa. Neste e em outros aspectos da presente invenção considerando um sinal de ruído de interferência, pressupõe-se que haja um acesso ao sinal de áudio por si só e ao sinal de interferência secundário por si só.

[0016] Em um outro aspecto, a presente invenção provê o controle da intensidade acústica específica de um sinal de áudio por meio da modificação do sinal de áudio, a fim de reduzir a diferença entre a sua intensidade acústica específica e uma intensidade acústica alvo específica.

[0017] Quando a intensidade acústica alvo específica não é uma função do sinal de áudio, a mesma pode ser uma intensidade acústica alvo específica armazenada ou recebida. Quando a intensidade acústica alvo específica não é uma função do sinal de áudio, a modificação ou a obtenção poderá explícita ou implicitamente calcular uma intensidade acústica específica ou uma intensidade acústica parcial específica. Exemplos de cálculo implícito incluem uma tabela de pesquisa ou

uma expressão matemática "de sentença fechada", na qual a intensidade acústica específica e/ou a intensidade acústica parcial específica é inerentemente determinada (o termo sentença fechada descreve uma expressão matemática que pode ser representada exatamente usando um número finito de operações e funções matemáticas padrão, tais como a exponencial ou a cosseno). Ainda, quanto a intensidade acústica alvo específica não é uma função do sinal de áudio, a intensidade acústica alvo específica pode ser tanto invariante no tempo e invariante na frequência, como também pode ser invariante no tempo.

[0018] Ainda, em um outro aspecto, a presente invenção provê o processamento de um sinal de áudio por meio do processamento do sinal de áudio ou uma medida do sinal de áudio de acordo com um ou mais processos e um ou mais parâmetros de controle de processo a fim de produzir uma intensidade acústica alvo específica. Embora a intensidade acústica alvo específica possa ser invariante no tempo ("fixa"), a intensidade acústica alvo específica poderá com vantagem ser uma função da intensidade acústica específica do sinal de áudio. Embora possa ser um sinal invariante no tempo e na frequência, tipicamente, o sinal de áudio em si é variante na frequência e no tempo, fazendo com que a intensidade acústica alvo específica seja variante na frequência e no tempo quando a mesma é uma função do sinal de áudio.

[0019] O áudio e uma intensidade acústica alvo específica ou uma representação de uma intensidade acústica alvo específica podem ser recebidos a partir de uma transmissão ou reproduzidos a partir de um meio de armazenamento.

[0020] A representação de uma intensidade acústica alvo específica pode ser um ou mais fatores de escala que gradua o sinal de áudio ou mede o sinal de áudio.

[0021] A intensidade acústica alvo específica de qualquer um dos

aspectos acima da presente invenção pode ser uma função do sinal de áudio ou medida do sinal de áudio. Uma medida adequada do sinal de áudio é a intensidade acústica específica do sinal de áudio. A função do sinal de áudio ou medida do sinal de áudio pode ser uma graduação do sinal de áudio ou medida do sinal de áudio. Por exemplo, a graduação pode ser uma ou uma combinação de graduações:

(a) uma graduação do fator de escala variante no tempo e variante na frequência $\Xi [b, t]$ da intensidade acústica específica, como na relação:

$$\hat{N}[b, t] = \Xi [b, t] N[b, t];$$

(b) uma graduação de fator de escala variante no tempo e invariante na frequência $\Phi [t]$ da intensidade acústica específica, como na relação:

$$\hat{N}[b, t] = \Phi [t] N[b, t];$$

(c) uma graduação do fator de escala invariante no tempo e variante na frequência $\Phi [b]$ da intensidade acústica específica, como na relação:

$$\hat{N}[b, t] = \Phi [b] N[b, t]; \text{ e}$$

(d) uma graduação do fator de escala invariante no tempo e invariante na frequência α da intensidade acústica específica, como na relação:

$$\hat{N}[b, t] = \alpha N[b, t],$$

na qual $\hat{N}[b, t]$ é a intensidade acústica alvo específica, $N[b, t]$ é a intensidade acústica específica do sinal de áudio, b é uma medida de frequência, e t é uma medida de tempo.

[0022] No caso (a) de um fator de escala variante no tempo e variante na frequência, a graduação pode ser determinada pelo menos em parte por uma razão de uma intensidade acústica multibanda desejada e de uma intensidade acústica multibanda do sinal de áudio. Tal graduação pode ser utilizável como um controle de faixa dinâmica. Outros

detalhes dos aspectos empregados pela presente invenção quanto a um controle de faixa dinâmica são apresentados a seguir.

[0023] Ainda no caso (a) de um fator de escala variante no tempo e na frequência, a intensidade acústica específica pode ser graduada por meio da razão de uma medida de uma forma espectral desejada para a medida de uma forma espectral do sinal de áudio. Esta graduação pode ser empregada no sentido de transformar o espectro percebido do sinal de áudio a partir de um espectro percebido variante no tempo para um espectro percebido invariante no tempo. Quando a intensidade acústica específica é graduada pela razão de uma medida de uma forma de espectro desejada para a medida de uma forma espectral do sinal de áudio, tal graduação pode ser utilizável como um equalizador dinâmico. Outros detalhes de se usar aspectos da presente invenção como um equalizador dinâmico são apresentados a seguir.

[0024] No caso (b) de um fator de escala variante no tempo e invariante na frequência, a graduação pode ser determinada pelo menos em parte por uma razão de uma intensidade acústica de banda larga desejada para a intensidade acústica de banda larga do sinal de áudio. Esta graduação pode ser utilizável como um controle de ganho automático ou como um controle de faixa dinâmica. Outros detalhes de se empregar aspectos da presente invenção como um controle de ganho automático ou controle de faixa dinâmica são apresentados a seguir.

[0025] No caso (a) (um fator de escala variante no tempo e na frequência) ou no caso (b) (um fator de escala variante no tempo e invariante na frequência), o fator de escala pode ser uma função do sinal de áudio ou medida do sinal de áudio.

[0026] No caso (c) de um fator de escala invariante no tempo, variante na frequência ou o caso (d) um fator de escala invariante no tempo e invariante na frequência, a modificação ou a obtenção pode incluir o armazenamento do fator de escala ou o fator de escala pode

ser recebido a partir de uma fonte externa.

[0027] Em ambos os casos (c) e (d), o fator de escala pode não ser uma função do sinal de áudio ou medida do sinal de áudio.

[0028] Em qualquer dos diversos aspectos da presente invenção e suas variações, a modificação, a obtenção, ou a produção podem, de várias maneiras, explícita ou implicitamente calcular (1) a intensidade acústica específica, e/ou (2) a intensidade acústica parcial específica, e/ou (3) a intensidade acústica alvo específica. Os cálculos implícitos podem envolver, por exemplo, uma tabela de pesquisa ou uma expressão matemática de sentença fechada.

[0029] Os parâmetros de modificação podem ser temporalmente suavizados. Os parâmetros de modificação podem ser, por exemplo, (1) uma pluralidade de fatores de graduação de amplitude relativos às bandas de frequência do sinal de áudio ou (2) uma pluralidade de coeficientes de filtro para o controle de um ou mais filtros, como, por exemplo, um filtro FIR de diversas derivações ou um filtro IIR multipolar. Os fatores de escala ou os coeficientes de filtro (e os filtros aos quais os mesmos são aplicados) podem ser variantes no tempo.

[0030] Ao calcular a função da intensidade acústica específica do sinal de áudio que define a intensidade acústica alvo específica ou o inverso desta função, o processo ou processos que fazem estes cálculos operam no que se pode ser caracterizado como o domínio de intensidade acústica perceptiva (psicoacústica) - a entrada e a saída do cálculo são as intensidades acústicas específicas. Em contrapartida, ao aplicar fatores de escala de amplitude às bandas de frequência do sinal de áudio ou ao aplicar coeficientes de filtro a uma filtragem controlável do sinal de áudio, os parâmetros de modificação operam no sentido de modificar o sinal de áudio fora do domínio de intensidade acústica perceptiva (psicoacústica) no que pode ser caracterizado como o domínio do sinal elétrico. Embora possam ser feitas modificações

ao sinal de áudio no domínio do sinal elétrico, estas mudanças no domínio do sinal elétrico são provenientes dos cálculos no domínio de intensidade acústica perceptiva (psicoacústica) de tal modo que o sinal de áudio modificado tenha uma intensidade acústica específica que se aproxima da intensidade acústica alvo específica desejada.

[0031] Ao se obter parâmetros de modificação a partir de cálculos no domínio da intensidade acústica, um controle maior sobre a intensidade acústica perceptiva e o equilíbrio espectral percebido poderá ser alcançado do que se estes parâmetros de modificação fossem obtidos no domínio do sinal elétrico. Além disso, o uso de um banco de filtros psicoacústico simulador de membrana basilar ou equivalente na realização dos cálculos de domínio de intensidade acústica pode prover um controle mais detalhado do espectro percebido que nas disposições que provêm parâmetros de modificação no domínio de sinal elétrico.

[0032] Cada modificação, obtenção e produção pode ser dependente de uma ou mais dentre uma medida de um sinal de áudio de interferência, uma intensidade acústica alvo específica, uma estimativa da intensidade acústica específica do sinal de áudio não modificado obtido a partir da intensidade acústica específica ou da intensidade acústica parcial específica do sinal de áudio modificado, a intensidade acústica específica do sinal de áudio não modificado, e uma aproximação à intensidade acústica alvo específica obtida a partir da intensidade acústica específica ou da intensidade acústica parcial específica do sinal de áudio modificado.

[0033] A modificação ou a obtenção pode produzir parâmetros de modificação pelo menos em parte a partir de um ou mais dentre uma medida de um sinal de áudio de interferência, uma intensidade acústica alvo específica, uma estimativa da intensidade acústica específica do sinal de áudio não modificado obtido a partir da intensidade acústi-

ca específica ou da intensidade acústica parcial específica do sinal de áudio modificado, a intensidade acústica específica do sinal de áudio não modificado, e uma aproximação à intensidade acústica alvo específica obtida a partir da intensidade acústica específica ou da intensidade acústica parcial específica do sinal de áudio modificado.

[0034] Mais particularmente, a modificação ou a obtenção pode produzir parâmetros de modificação pelo menos em parte a partir de:

(1) um dentre:

- uma intensidade acústica alvo específica; e
- uma estimativa da intensidade acústica específica do sinal

de áudio não modificado recebido a partir da intensidade acústica específica do sinal de áudio modificado, e

(2) um dentre:

- a intensidade acústica específica do sinal de áudio não modificado, e
- uma aproximação à intensidade acústica alvo específica obtida a partir da intensidade acústica específica do sinal de áudio modificado,

ou quando um sinal de áudio de interferência deve ser levado em consideração, a modificação ou a obtenção pode derivar de parâmetros de modificação pelo menos em parte de:

(1) uma medida de um sinal de áudio de interferência;

(2) um dentre:

- uma intensidade acústica alvo específica, e
- uma estimativa da intensidade acústica específica do sinal de áudio não modificado obtido a partir da intensidade acústica parcial específica do sinal de áudio modificado, e

(3) um dentre:

- a intensidade acústica específica do sinal de áudio não modificado, e

- uma aproximação à intensidade acústica alvo específica obtida a partir da intensidade acústica parcial específica do sinal de áudio modificado.

[0035] Uma disposição de alimentação direta pode ser empregada na qual a intensidade acústica específica é obtida a partir do sinal de áudio, e na qual a intensidade acústica alvo específica é recebida a partir de uma fonte externa ao método ou a partir de um armazenamento quando a modificação ou a obtenção inclui o armazenamento de uma intensidade acústica alvo específica. De maneira alternativa, uma disposição híbrida de alimentação direta / retroalimentação pode ser empregada, na qual uma aproximação à intensidade acústica alvo específica é obtida a partir do sinal de áudio modificado, e na qual a intensidade acústica alvo específica é recebida de uma fonte externa ao método ou a partir de um armazenamento quando a modificação ou a obtenção inclui o armazenamento de uma intensidade acústica alvo específica.

[0036] A modificação ou a obtenção pode incluir um ou mais processos para a obtenção, explícita ou implícita, da intensidade acústica alvo específica, um ou alguns dos quais calcula, explícita ou implicitamente, uma função do sinal de áudio ou a medida do sinal de áudio. Em uma alternativa, uma disposição de alimentação direta pode ser empregada na qual a intensidade acústica específica e a intensidade acústica alvo específica são obtidas a partir do sinal de áudio, a obtenção da intensidade acústica alvo específica empregando uma função do sinal de áudio ou a medida do sinal de áudio. Em uma outra alternativa, uma disposição híbrida de alimentação direta / retroalimentação pode ser empregada, na qual uma aproximação à intensidade acústica alvo específica é obtida a partir do sinal de áudio modificado, e a intensidade acústica alvo específica é obtida a partir do sinal de áudio, a obtenção da intensidade acústica alvo específica empregando

uma função do sinal de áudio ou a medida do sinal de áudio.

[0037] A modificação ou a obtenção pode incluir um ou mais processos para a obtenção, explícita ou implícita, de uma estimativa da intensidade acústica específica do sinal de áudio não modificado em resposta ao sinal de áudio modificado, um ou alguns dos quais calcula, explícita ou implicitamente, o inverso de uma função do sinal de áudio ou a medida do sinal de áudio. Em uma alternativa, uma disposição de retroalimentação é empregada na qual uma estimativa da intensidade acústica específica do sinal de áudio não modificado e uma aproximação à intensidade acústica alvo específica são obtidas a partir do sinal de áudio modificado, a estimativa da intensidade acústica específica sendo calculada usando o inverso da dita função do sinal de áudio ou da medida do sinal de áudio. Em uma outra alternativa, uma disposição híbrida de alimentação direta / retroalimentação é empregada na qual a intensidade acústica específica é obtida a partir do sinal de áudio, e a estimativa da intensidade acústica específica do sinal de áudio não modificado é obtida a partir do sinal de áudio modificado, a obtenção da estimativa sendo calculada usando o inverso da dita função do sinal de áudio ou da medida do sinal de áudio.

[0038] Os parâmetros de modificação podem ser aplicados ao sinal de áudio de modo a produzir um sinal de áudio modificado.

[0039] Um outro aspecto da presente invenção é que pode haver uma separação temporal e/ou espacial de processos ou dispositivos de modo a existir, com efeito, um codificador ou codificação e ainda um decodificador ou decodificação. Por exemplo, pode haver um sistema de codificação / decodificação no qual a modificação ou obtenção pode transmitir ou receber ou armazenar ou ainda reproduzir o sinal de áudio e ainda (1) os parâmetros de modificação ou (2) uma intensidade acústica alvo específica ou uma representação de uma intensidade acústica alvo específica. De maneira alternativa, pode ha-

ver, com efeito, somente um codificador ou codificação na qual se faz uma transmissão ou armazenamento do sinal de áudio e (1) dos parâmetros de modificação ou (2) de uma intensidade acústica alvo específica ou representação da intensidade acústica alvo específica. De maneira alternativa, conforme acima mencionado, pode haver, com efeito, somente um decodificador ou decodificação na qual se faz uma recepção ou reprodução do sinal de áudio e (1) dos parâmetros de modificação ou (2) de uma intensidade acústica alvo específica ou representação da intensidade acústica alvo específica.

DESCRIÇÃO DOS DESENHOS

[0040] A Figura 1 é um diagrama em bloco funcional ilustrando um exemplo de uma implementação de alimentação direta de acordo com os aspectos da presente invenção.

[0041] A Figura 2 é um diagrama em bloco funcional ilustrando um exemplo de uma implementação de retroalimentação de acordo com os aspectos da presente invenção.

[0042] A Figura 3 é um diagrama em bloco funcional ilustrando um exemplo de uma implementação híbrida de alimentação direta / retroalimentação de acordo com os aspectos da presente invenção.

[0043] A Figura 4 é um diagrama em bloco funcional ilustrando um exemplo de uma outra implementação híbrida de alimentação direta / retroalimentação de acordo com os aspectos da presente invenção.

[0044] A Figura 5 é um diagrama em bloco funcional ilustrando a maneira pela qual o sinal de áudio não modificado e os parâmetros de modificação conforme determinados por qualquer uma dentre as disposições de alimentação direta, de retroalimentação, ou híbridas de alimentação direta / retroalimentação podem ser armazenados ou transmitidos para uso, por exemplo, em um dispositivo ou processo temporal ou espacialmente separado.

[0045] A Figura 6 é um diagrama em bloco funcional ilustrando a

maneira pela qual o sinal de áudio não modificado e a intensidade acústica alvo específica ou representação da mesma conforme determinados por qualquer uma dentre as disposições de alimentação direta, de retroalimentação, ou híbridas de alimentação direta / retroalimentação podem ser armazenados ou transmitidos para uso, por exemplo, em um dispositivo ou processo temporal ou espacialmente separado.

[0046] A Figura 7 é um diagrama em bloco funcional esquemático ou fluxograma esquemático mostrando um resumo de um aspecto da presente invenção.

[0047] A Figura 8 é uma resposta característica idealizada de um filtro linear $P(z)$ adequado como um filtro de transmissão em uma modalidade da presente invenção na qual o eixo vertical é a atenuação em decibéis (dB) e o eixo horizontal é uma base logarítmica 10 de frequência energia Hertz (Hz).

[0048] A Figura 9 mostra a relação entre a escala de frequência ERB (eixo vertical) e a frequência em Hertz (eixo horizontal).

[0049] A Figura 10 mostra um conjunto de respostas características de filtro auditivo idealizadas que aproximam uma formação de banda crítica na escala ERB. A escala horizontal é a frequência em Hertz e a escala vertical é o nível em decibéis.

[0050] A Figura 11 mostra os contornos de intensidade acústica iguais do ISO 226. A escala horizontal é a frequência em Hertz (escala de base 10 logarítmica) e a escala vertical é o nível de pressão de som em decibéis.

[0051] A Figura 12 mostra os contornos de intensidade acústica iguais do ISO 226 normalizados pelo filtro de transmissão $P(z)$. A escala horizontal é a frequência em Hertz (escala de base 10 logarítmica) e a escala vertical é o nível de pressão de som em decibéis.

[0052] A Figura 13a é um gráfico idealizado mostrando os ganhos

de banda larga e os ganhos de multibanda para uma graduação de intensidade acústica de 0,25 em um segmento de fala feminina. A escala horizontal é de bandas ERB e a escala vertical é o ganho relativo em decibéis (dB).

[0053] A Figura 13b é um gráfico idealizado mostrando a intensidade acústica específica, respectivamente, de um sinal original, de um sinal modificado em ganho de banda larga, e um sinal modificado em ganho de multibanda. A escala horizontal é de bandas ERB e a escala vertical é de uma intensidade acústica específica (sone/ERB).

[0054] A Figura 14a é um gráfico idealizado mostrando: $L_o [t]$ como uma função de $L_1 [t]$ para um ganho ACG típico. A escala horizontal é o $\log (L_1 [t])$ e a escala vertical é o $\log (L_o [t])$.

[0055] A Figura 14b é um gráfico idealizado mostrando: $L_o [t]$ como uma função de $L_1 [t]$ para um controle DRC típico. A escala horizontal é o $\log (L_1 [t])$ e a escala vertical é o $\log (L_o [t])$.

[0056] A Figura 15 é um gráfico idealizado mostrando uma função típica de suavização de banda para um controle DRC multibanda. A escala horizontal é um número de banda e a escala vertical é a saída de ganho para a banda b .

[0057] A Figura 16 é um diagrama em bloco funcional esquemático ou fluxograma esquemático mostrando um resumo de um aspecto da presente invenção.

[0058] A Figura 17 é um diagrama em bloco funcional esquemático ou fluxograma esquemático similar ao da Figura 1 incluindo uma compensação para ruído em um ambiente de reprodução.

MELHOR MODO DE SE EXECUTAR A INVENÇÃO

[0059] As Figuras 1 a 4 mostram diagramas em blocos funcionais ilustrando as possíveis alimentação direta, a retroalimentação, e duas versões de exemplos de implementação híbrida de alimentação direta / retroalimentação de acordo com aspectos da presente invenção.

[0060] Com referência ao exemplo de uma topologia de alimentação direta na Figura 1, um sinal de áudio é aplicado em dois trajetos: (1) um trajeto de sinal tendo um processo ou dispositivo 2 ("Modificar Sinal de áudio") capaz de modificar o áudio em resposta aos parâmetros de modificação, e (2) um trajeto de controle tendo um processo ou dispositivo 4 ("Gerar Parâmetros de Modificação") capaz de gerar estes parâmetros de modificação. O exemplo de topologia de alimentação direta Modificar Sinal de áudio 2 na Figura 1 e em cada um dos exemplos das Figuras 2 a 4 pode ser um dispositivo ou processo que modifica o sinal de áudio, por exemplo, sua amplitude, de uma maneira variante na frequência e/ou no tempo de acordo com os parâmetros de modificação M recebidos do Gerar Parâmetros de Modificação 4 (ou dos processos ou dispositivos de contrapartida 4', 4'', e 4''' em cada uma dos exemplos das Figuras 2 a 4, respectivamente). O Gerar Parâmetros de Modificação 4 e suas contrapartidas nas Figuras 2 a 4 opera pelo menos parcialmente no domínio de intensidade acústica perceptiva. O Modificar Sinal de áudio 2 opera no domínio do sinal elétrico e produz um sinal de áudio modificado em cada um dos exemplos das Figuras 1 a 4. Ainda nos exemplos das Figuras 1 a 4, o Modificar Sinal de áudio 2 e o Gerar Parâmetros de Modificação 4 (ou suas contrapartidas) modificam o sinal de áudio de modo a reduzir a diferença entre a intensidade acústica específica e uma intensidade acústica alvo específica.

[0061] No exemplo de alimentação direta da Figura 1, o processo ou dispositivo 4 pode incluir vários processos e/ou dispositivos; um processo ou dispositivo "Calcular Intensidade acústica alvo específica" 6 que calcula uma intensidade acústica alvo específica em resposta ao sinal de áudio ou uma medida de sinal de áudio, tal como a intensidade acústica específica do sinal de áudio, um processo ou dispositivo "Calcular Intensidade acústica específica" 8 que calcula a intensidade

acústica específica do sinal de áudio em resposta ao sinal de áudio ou uma medida dos sinais de áudio tais como sua excitação, e um processo ou dispositivo "Calcular Parâmetros de Modificação" 10 que calcula os parâmetros de modificação em resposta à intensidade acústica específica e à intensidade acústica alvo específica. O Calcular Intensidade acústica alvo específica 6 pode realizar uma ou mais funções "F", cada uma das quais podendo ter parâmetros de função. Por exemplo, ele pode calcular a intensidade acústica específica do sinal de áudio e em seguida aplicar uma ou mais funções F ao mesmo de modo a prover uma intensidade acústica alvo específica. Isto é indicado esquematicamente na Figura 1 como uma entrada "Selecionar Funções F e Parâmetros de Funções" para o processo ou dispositivo 6. Ao invés de ser calculado pelo dispositivo ou processo 6, a intensidade acústica alvo específica pode ser provida por um processo ou dispositivo de armazenamento (mostrado esquematicamente como uma entrada "Armazenado" para o processo ou dispositivo 10) incluído ou associado ao Gerar Parâmetros de Modificação 4, ou por uma fonte externa ao processo ou dispositivo como um todo (mostrada esquematicamente como a entrada "Externa" ao processo ou dispositivo 10). Deste modo, os parâmetros de modificação se baseiam pelo menos em parte nos cálculos do domínio da intensidade acústica perceptiva (psicoacústico) (isto é, pelo menos a intensidade acústica específica e, em alguns casos, os cálculos de intensidade acústica alvo específica).

[0062] Os cálculos feitos pelos processos ou dispositivos 6, 8 (e pelos processos ou dispositivos 12, 14, 10' da Figura 2, 6, 14, 10" da Figura 3, e 8, 12, 10"" da Figura 4) podem ser feitos explícita ou implicitamente. Exemplos de um desempenho implícito incluem (1) uma tabela de pesquisa cujas entradas se baseiam no todo ou em parte na intensidade acústica específica e/ou na intensidade acústica alvo específica e/ou nos cálculos de parâmetro de modificação, e (2) uma ex-

pressão matemática de sentença fechada inerentemente baseada no todo ou em parte na intensidade acústica específica e/ou na intensidade acústica alvo específica e/ou nos parâmetros de modificação.

[0063] Embora os processos ou dispositivos de cálculo 6, 8, 10 da Figura 1 (e os processos ou dispositivos 12, 14, 10' da Figura 2, 6, 14, 10" da Figura 3, e 8, 12, 10'" da Figura 4) sejam mostrados esquematicamente e descritos em separado, isto é, apenas para fins de explicação. Será entendido que alguns ou todos estes processos ou dispositivos podem ser combinados em um único processo ou dispositivo ou combinados de várias maneiras em múltiplos processos ou dispositivos. Por exemplo, na disposição da Figura 9 abaixo, uma topologia de alimentação direta como no exemplo da Figura 1, o processo ou dispositivo que calcula os parâmetros de modificação atua desta forma em resposta à excitação suavizada obtida a partir do sinal de áudio e de uma intensidade acústica alvo específica. Na Figura 9, o dispositivo ou processo que calcula os parâmetros de modificação implicitamente calcula a intensidade acústica específica do sinal de áudio.

[0064] Como um aspecto da presente invenção, na Figura 1 e em outros exemplos de modalidades da presente invenção, a intensidade acústica alvo específica ($N[b, t]$) pode ser calculada por meio da graduação da intensidade acústica específica ($N[b, t]$) com um ou mais fatores de graduação. A graduação pode ser uma graduação de fator de escala variante no tempo e na frequência $\Xi [b, t]$ da intensidade acústica específica como na relação

$$\hat{N}[b, t] = \Xi [b, t] N[b, t],$$

uma graduação de fator de escala variante no tempo, invariante na frequência $\Phi [t]$ da intensidade acústica específica como na relação

$$\hat{N}[b, t] = \Phi [t] N[b, t],$$

uma graduação de fator de escala invariante no tempo, variante na frequência $\Theta [b]$ da intensidade acústica específica como na relação

$$\hat{N}[b, t] = \Theta[b] N[b, t], \text{ ou}$$

uma graduação de fator de escala α da intensidade acústica específica como na relação

$$N[b, t] = \alpha N[b, t],$$

onde b é uma medida de frequência (por exemplo, o número da banda) e t é uma medida de tempo (por exemplo, o número do bloco). Múltiplas graduações podem também ser empregadas, usando múltiplas instâncias de uma graduação em particular e/ou combinações de graduações particulares. Exemplos destas múltiplas graduações são apresentados abaixo. Em alguns casos, conforme explicado em mais detalhe abaixo, a graduação pode ser uma função do sinal de áudio ou a medida do sinal de áudio. Em outros casos, também conforme explicado em mais detalhe abaixo, quando a graduação não é uma função de uma medida do sinal de áudio, a graduação poderá ser determinada ou suprida de outra maneira. Por exemplo, um usuário pode selecionar ou aplicar um fator de escala invariante no tempo e frequência α ou um fator de escala invariante no tempo, variante na frequência $\Theta[b]$.

[0065] Assim, a intensidade acústica alvo específica pode ser expressa como uma ou mais funções F do sinal de áudio ou medida do sinal de áudio (a intensidade acústica específica sendo uma medida possível do sinal de áudio):

$$\hat{N}[b, t] = F(N[b, t]).$$

[0066] Desde que a função ou funções F seja inversivo, a intensidade acústica específica ($N[b, t]$) do sinal de áudio não modificado poderá ser calculada como a função ou funções F^{-1} inversa da intensidade acústica alvo específica ($\hat{N}[b, t]$):

$$N[b, t] = F^{-1}(\hat{N}[b, t]).$$

[0067] Como se notará, a função ou funções F^{-1} inversa é calculada nos exemplos de alimentação e alimentação híbrida de alimentação direta / retroalimentação das Figuras 2 e 4.

[0068] Uma entrada "Selecionar Funções e Parâmetros de Funções" para o Calcular Intensidade acústica alvo específica 6 é mostrada para indicar que o dispositivo ou processo 6 pode calcular a intensidade acústica alvo específica por meio da aplicação de uma ou mais funções de acordo com um ou mais parâmetros de função. Por exemplo, o Calcular Intensidade acústica alvo específica 8 pode calcular a função ou funções "F" da intensidade acústica específica do sinal de áudio a fim de definir a intensidade acústica alvo específica. Por exemplo, a entrada "Selecionar Funções e Parâmetros de Função" pode selecionar uma ou mais funções em particular que recaia no um ou mais tipos acima de graduação, juntamente com um ou mais parâmetros de função, tais como as constantes (por exemplo, os fatores de escala) pertencentes às funções.

[0069] Os fatores de graduação associados a uma graduação podem servir como uma representação da intensidade acústica alvo específica, uma vez que a intensidade acústica alvo específica pode ser computada como uma graduação da intensidade acústica específica, conforme indicado acima. Deste modo, no exemplo da Figura 9, descrita abaixo e acima mencionada, a tabela de pesquisa pode ser indexada pelos fatores de escala e excitações, de tal modo que o cálculo da intensidade acústica específica e da intensidade acústica alvo específica fique inerente à tabela.

[0070] Ao se empregar uma tabela de pesquisa, uma expressão matemática de sentença fechada, ou alguma outra técnica, a operação do Gerar Parâmetros de Modificação 4 (e seus processos ou dispositivos similares 4', 4" e 4''' das Figuras 2 a 4) é tal que os cálculos se baseiem no domínio da intensidade acústica perceptiva (psicoacústico) mesmo que a intensidade acústica específica ou a intensidade acústica alvo específica não tenha sido explicitamente calculada. Em ambos os casos, há uma intensidade acústica específica explícita ou há uma

intensidade acústica específica nocional, implícita. De maneira similar, em ambos os casos, há uma intensidade acústica alvo específica explícita ou há uma intensidade acústica alvo específica nocional, implícita. Em qualquer caso, o cálculo dos parâmetros de modificação busca gerar parâmetros de modificação que modifiquem o sinal de áudio a fim de reduzir a diferença entre a intensidade acústica específica e uma intensidade acústica alvo específica.

[0071] Em um ambiente de reprodução tendo um sinal de áudio de interferência secundário, tal como um ruído, o Calcular Parâmetros de Modificação 10 (e seus processos ou dispositivos similares 10', 10", 10''' em cada uma das Figuras 2 a 4, respectivamente) pode também receber como uma entrada opcional uma medida de tal sinal de áudio de interferência secundário ou o próprio sinal de interferência secundário como uma de suas entradas. Tal entrada opcional é mostrada na Figura 1 (e nas Figuras 2 a 4) em linha pontilhada. A medida de um sinal de interferência secundário pode ser a sua excitação, como na Figura 17, descrita abaixo. A aplicação de uma medida do sinal de interferência ou do próprio sinal (pressupondo que o sinal de interferência esteja separadamente disponível para processamento) ao Calcular Parâmetros de Modificação, o processo ou dispositivos 10 da Figura 1 (ou seus processos ou dispositivos similares 10', 10" e 10''' de cada uma das Figuras 2 a 4, respectivamente) permite que tal processo ou dispositivo adequadamente configurado calcule os parâmetros de modificação que levam o sinal de interferência em consideração, conforme explicado abaixo no parágrafo "Compensação de Ruído". Nos exemplos das Figuras 2 a 4, o cálculo da intensidade acústica parcial específica pressupõe que uma medida adequada de um sinal de interferência é aplicado não somente ao respectivo Calcular Parâmetros de Modificação 10', 10", 10''', mas também ao processo ou dispositivo 12 "Calcular Aproximação de Intensidade acústica específica de Áudio

Não Modificado" e/ou ao processo ou dispositivo 14 "Calcular Aproximação de Intensidade acústica alvo específica" a fim de facilitar o cálculo da intensidade acústica parcial específica por aquela função ou dispositivo. No exemplo de alimentação direta da Figura 1, a intensidade acústica parcial específica não é explicitamente calculada - o Calcular Parâmetros de Modificação 10 da Figura 1 calcula os parâmetros de modificação apropriados para que a intensidade acústica parcial específica do áudio modificado se aproxime da intensidade acústica alvo específica. Isto é explicado em mais detalhe abaixo no parágrafo "Compensação de Ruído", acima mencionado.

[0072] Conforme acima mencionado, em cada um dos exemplos das Figuras 1 a 4, os parâmetros de modificação M, quando aplicados ao sinal de áudio pelo Modificador de Sinal de áudio 2, reduzem a diferença entre a intensidade acústica específica ou a intensidade acústica parcial específica do áudio modificado resultante e a intensidade acústica alvo específica. Em termos ideais, a intensidade acústica específica do sinal de áudio modificado se aproxima muito ou é igual à intensidade acústica alvo específica. Os parâmetros de modificação M, por exemplo, assumem a forma de fatores de ganho variados aplicados às bandas de frequência obtidas a partir de um banco de filtros ou aos coeficientes de um filtro variante no tempo. Por conseguinte, em todos os exemplos das Figuras 1 a 4, o Modificar Sinal de áudio 2 pode ser implementado, por exemplo, como uma pluralidade de escaladores de amplitude, cada qual operando em uma banda de frequência, ou em um filtro variante no tempo (por exemplo, um filtro FIR de diversas derivações ou um filtro IIR multipolar).

[0073] Aqui ou em qualquer ponto no presente documento, o uso do mesmo numeral de referência indica que o dispositivo ou processo pode ser substancialmente idêntico ao outro ou a outros que tenham o mesmo numeral de referência. Os numerais de referência com núme-

ros aspeados (por exemplo, 10'') indicam que o dispositivo ou processo é similar em estrutura ou função, mas pode ser uma modificação de um outro ou outros que tenham o mesmo numeral de referência básico ou suas versões iniciadas.

[0074] Sob certas limitações, uma disposição de retroalimentação quase equivalente ao exemplo de alimentação direta da Figura 1 pode ser produzida. A Figura 2 ilustra tal exemplo no qual o sinal de áudio é também aplicado a um processo ou dispositivo 2 Modificar Sinal de áudio em um trajeto de sinal. O processo ou dispositivo 2 também recebe os parâmetros de modificação M de um trajeto de controle no qual um processo ou dispositivo 4' Gerar Parâmetros de Modificação em um dispositivo de retroalimentação recebe como entrada o sinal de áudio modificado a partir da saída do Modificar Sinal de áudio 2. Desta forma, na Figura 2, o áudio modificado, ao invés do áudio não modificado, é aplicado ao trajeto de controle. O processo ou dispositivo 2 Modificar Sinal de áudio e o processo ou dispositivo 4' Gerar Parâmetros de Modificação modificam o sinal de áudio a fim de reduzir a diferença entre a sua intensidade acústica específica e uma intensidade acústica alvo específica. O processo ou dispositivo 4 pode incluir várias funções e ou dispositivos: um processo ou dispositivo 12 "Calcular Aproximação de Intensidade acústica específica de Áudio Não Modificado", um processo ou dispositivo 14 "Calcular Aproximação de Intensidade acústica alvo específica", e um processo ou dispositivo 10' "Calcular Parâmetros de Modificação" que calcula os parâmetros de modificação.

[0075] Com a limitação de que a função ou funções F é inversiva, o processo ou dispositivo 12 estima a intensidade acústica específica do sinal de áudio não modificado ao aplicar a função inversa F^{-1} à intensidade acústica específica ou à intensidade acústica parcial específica do sinal de áudio modificado. O dispositivo ou processo 12 pode

calcular uma função inversa F^{-1} conforme acima descrito. Isto é indicado esquematicamente na Figura 2 como uma entrada "Selecionar Funções F^{-1} Inversas e Parâmetros de Função" para o processo ou dispositivo 12. O "Calcular Aproximação de Intensidade acústica alvo específica" 14 opera por meio do cálculo da intensidade acústica específica ou intensidade acústica parcial específica do sinal de áudio modificado. Tal intensidade acústica ou intensidade acústica parcial específica é uma aproximação da intensidade acústica alvo específica. A aproximação da intensidade acústica específica do sinal de áudio não modificado e a aproximação da intensidade acústica alvo específica são usadas pelo Calcular Parâmetros de Modificação 10' a fim de obter os parâmetros de modificação M, que, quando aplicado ao sinal de áudio pelo Modificar Sinal de áudio 2, reduz a diferença entre a intensidade acústica específica ou a intensidade acústica parcial específica do áudio modificado e a intensidade acústica alvo específica. Conforme acima mencionado, estes parâmetros de modificação M podem, por exemplo, assumir a forma de ganhos variantes no tempo aplicados às bandas de frequência de um banco de filtros ou aos coeficientes de um filtro variante no tempo. Nas modalidades práticas do Calcular Parâmetros de Modificação 10", um laço de retroalimentação poderá introduzir um retardo entre a computação e a aplicação dos parâmetros de modificação M.

[0076] Conforme acima mencionado, em um ambiente de reprodução tendo um sinal de áudio de interferência secundário, tal como um ruído, o Calcular Parâmetros de Modificação 10', o Calcular Aproximação de Intensidade acústica específica de Áudio Não Modificado 12, e o Calcular Aproximação da Intensidade acústica alvo específica 14 poderão também receber como uma entrada opcional uma medida de tal sinal de áudio de interferência secundário ou o próprio sinal de interferência secundário como uma de suas entradas e o processo ou

dispositivo 12 ou o processo ou dispositivo 14 poderão calcular a intensidade acústica parcial específica do sinal de áudio modificado. Estas entradas opcionais são mostradas na Figura 2, nas linhas pontilhadas.

[0077] Conforme acima mencionado, exemplos de implementação híbrida de alimentação direta / retroalimentação dos aspectos da presente invenção são possíveis. As Figuras 3 e 4 mostram dois exemplos de tais implementações. Nas Figuras 3 e 4, assim como nas Figuras 1 e 2, um sinal de áudio é também aplicado a um processo ou dispositivo Modificar Sinal de áudio em um trajeto de sinal, mas o Gerar Parâmetros de Modificação (4" na Figura 3 e 4'" na Figura 4) em respectivos trajetos de controle recebe tanto um sinal de áudio não modificado como um sinal de áudio modificado. Em ambas as Figuras 3 e 4, o Modificar Sinal de áudio 2 e o Gerar Parâmetros de Modificação (4" e 4'", respectivamente) modificam o sinal de áudio a fim de reduzir a diferença entre sua intensidade acústica específica, que pode ser implícita, e uma intensidade acústica alvo específica, que também pode ser implícita.

[0078] No exemplo da Figura 3, o processo ou dispositivo 4' Gerar Parâmetros de Modificação pode incluir diversas funções e ou dispositivos: um Calcular Intensidade acústica alvo específica 6, como na Figura 1, um Calcular Aproximação de Intensidade acústica alvo específica 14, como na retroalimentação da Figura 2, e um processo ou dispositivo 10" "Calcular Parâmetros de Modificação". Como na Figura 1, na porção de alimentação direta deste exemplo híbrido de alimentação direta / retroalimentação, o Calcular Intensidade acústica alvo específica 6 pode realizar uma ou mais funções "F", cada uma das quais podendo ter parâmetros de função. Isto é indicado esquematicamente na Figura 3 como uma entrada "Selecionar Funções F e Parâmetros de Função" ao processo ou dispositivo 6. Na porção de retroalimentação

deste exemplo híbrido de alimentação direta / retroalimentação, o sinal de áudio modificado é aplicado a um Calcular Aproximação de Intensidade acústica alvo específica 14, como na retroalimentação da Figura 2. O processo ou dispositivo 14 opera no exemplo da Figura 3 e no exemplo da Figura 2 por meio do cálculo da intensidade acústica específica ou da intensidade acústica parcial específica do sinal de áudio modificado. Tal intensidade acústica específica ou intensidade acústica parcial específica é uma aproximação da intensidade acústica alvo específica. A intensidade acústica alvo específica (do processo ou dispositivo 6) e a aproximação da intensidade acústica alvo específica (do processo ou dispositivo 14) são aplicadas ao Calcular Parâmetros de Modificação 10" a fim de obter os parâmetros de modificação M que, quando aplicados ao sinal de áudio pelo Modificar Sinal de áudio 2, reduzem a diferença entre a intensidade acústica específica do sinal de áudio não modificado e a intensidade acústica alvo específica. Conforme acima mencionado, os parâmetros de modificação M, podem, por exemplo, assumir a forma de ganhos variantes no tempo aplicados às bandas de frequência de um banco de filtros ou aos coeficientes de um filtro variante no tempo. Nas modalidades práticas, um laço de retroalimentação poderá introduzir um retardo entre a computação e a aplicação dos parâmetros de modificação M. Conforme acima mencionado, em um ambiente de reprodução tendo um sinal de áudio de interferência secundário, tal como um ruído, o Calcular Parâmetros de Modificação 10", e o Calcular Aproximação de Intensidade acústica alvo específica 14 poderão também receber como uma entrada opcional uma medida de tal sinal de áudio de interferência secundário ou o próprio sinal de interferência secundário como uma de suas entradas e o processo ou dispositivo 14 poderão calcular a intensidade acústica parcial específica do sinal de áudio modificado. Estas entradas opcionais são mostradas na Figura 3, nas linhas pontilhadas.

[0079] O Calcular Parâmetros de Modificação 10" poderá empregar um dispositivo ou função de detecção de erro, de tal modo que as diferenças entre a sua intensidade acústica alvo específica e as entradas de aproximação de intensidade acústica alvo de especificação ajustem os Parâmetros de Modificação de modo a reduzir as diferenças entre a aproximação da intensidade acústica alvo específica e a intensidade acústica alvo específica "corrente". Estes ajustes reduzem as diferenças entre a intensidade acústica específica do sinal de áudio não modificado e a intensidade acústica alvo específica, que pode ser implícita. Desta forma, os parâmetros de modificação M podem ser atualizados com base em um erro entre a intensidade acústica alvo específica, computada no trajeto da alimentação direta a partir da intensidade acústica específica do áudio original usando a função F, e a aproximação da intensidade acústica alvo específica computada no trajeto de retroalimentação a partir da intensidade acústica específica ou da intensidade acústica parcial específica do áudio modificado.

[0080] Na Figura 4, é mostrado um exemplo de alimentação direta / retroalimentação alternativa. Esta alternativa difere do exemplo da Figura 3 no sentido de que a função ou funções F^{-1} inversa é calculada no trajeto de retroalimentação ao invés de a função ou funções F ser calculada no trajeto de alimentação direta. Na Figura 4, o processo ou dispositivo 4' Gerar Parâmetros de Modificação pode incluir diversas funções e ou dispositivos: um Calcular Intensidade acústica específica 8, como no exemplo da alimentação direta da Figura 1, um Calcular Aproximação de Intensidade acústica específica de Áudio Não Modificado 12, como no exemplo da retroalimentação da Figura 2, e um Calcular Parâmetros de Modificação 10". Como no exemplo da alimentação direta da Figura 1, o Calcular Intensidade acústica específica 8 provê, como uma entrada ao Calcular Parâmetros de Modificação 10", a intensidade acústica específica do sinal de áudio não modificado.

Como no exemplo de retroalimentação da Figura 2, com a limitação de que a função ou funções F ser inversiva, o processo ou dispositivo 12 estima a intensidade acústica específica do sinal de áudio não modificado por meio da aplicação da função F^{-1} inversa à intensidade acústica específica ou à intensidade acústica parcial específica do sinal de áudio modificado. Uma entrada "Selecionar Funções Inversas e Parâmetros de Funções Inversas" para o Calcular Aproximação de Intensidade acústica específica de Áudio Não Modificado 12 é mostrada para indicar que o dispositivo ou processo 12 pode calcular uma função inversa F^{-1} , conforme acima descrito. Isto é indicado esquematicamente na Figura 4 como uma entrada "Selecionar Funções Inversas F^{-1} e Parâmetros de Função" ao processo ou dispositivo 12. Desta forma, o pedido 12 provê como uma outra entrada ao Calcular Parâmetros de Modificação 10" uma aproximação à intensidade acústica específica do sinal de áudio não modificado.

[0081] Como nos exemplos das Figuras 1 a 3, o Calcular Parâmetros de Modificação 10" obtém os parâmetros de modificação M que, quando aplicados ao sinal de áudio pelo Modificar Sinal de áudio 2, reduzem a diferença entre a intensidade acústica específica do sinal de áudio não modificado e a intensidade acústica alvo específica, que é implícita neste exemplo. Conforme acima mencionado, os parâmetros de modificação M , podem, por exemplo, assumir a forma de ganhos variantes no tempo aplicados às bandas de frequência de um banco de filtros ou aos coeficientes de um filtro variante no tempo. Nas modalidades práticas, um laço de retroalimentação poderá introduzir um retardo entre a computação e a aplicação dos parâmetros de modificação M . Conforme acima mencionado, em um ambiente de reprodução tendo um sinal de áudio de interferência secundário, tal como um ruído, o Calcular Parâmetros de Modificação 10" e o Calcular Aproximação de Intensidade acústica específica do Áudio Não Modificado 12

poderão também receber como uma entrada opcional uma medida de tal sinal de áudio de interferência secundário ou o próprio sinal de interferência secundário como uma de suas entradas e o processo ou dispositivo 12 poderá calcular a intensidade acústica parcial específica do sinal de áudio modificado. Estas entradas opcionais são mostradas na Figura 4, nas linhas pontilhadas.

[0082] O Calcular Parâmetros de Modificação 10" pode empregar um dispositivo ou função de detecção de erro, de tal modo que as diferenças entre a sua intensidade acústica específica e as entradas de aproximação de intensidade acústica específica produzam saídas que ajustem os Parâmetros de Modificação de modo a reduzir as diferenças entre a aproximação da intensidade acústica específica e a intensidade acústica específica "corrente". Uma vez que a aproximação da intensidade acústica específica é obtida a partir da intensidade acústica específica ou da intensidade acústica parcial específica do áudio modificado, que pode ser vista como uma aproximação da intensidade acústica alvo específica, estes ajustes reduzem as diferenças entre a intensidade acústica específica do sinal de áudio modificado e a intensidade acústica alvo específica, que é inerente na função ou funções F^{-1} . Desta forma, os parâmetros de modificação M podem ser atualizados com base em um erro entre a intensidade acústica específica, computada no trajeto da alimentação direta do áudio original, e a aproximação da intensidade acústica específica computada, usando a função ou funções F^{-1} inversa, no trajeto de retroalimentação a partir da intensidade acústica específica ou da intensidade acústica parcial específica do áudio modificado. Devido ao trajeto de retroalimentação, implementações práticas poderão introduzir um retardo entre a atualização e a aplicação dos parâmetros de modificação.

[0083] Embora os parâmetros de modificação M nos exemplos das Figuras 1 a 4, quando aplicados a um processo ou dispositivo 2 Modi-

ficar Sinal de áudio, reduzam a diferença entre a intensidade acústica específica do sinal de áudio e a intensidade acústica alvo específica, nas modalidades práticas, os parâmetros de modificação correspondentes produzidos em resposta ao mesmo sinal de áudio poderão não ser idênticos entre si.

[0084] Embora não crítico ou essencial aos aspectos da presente invenção, o cálculo da intensidade acústica específica do sinal de áudio ou do sinal de áudio modificado poderá com vantagem empregar as técnicas apresentadas no dito Pedido de Patente Internacional N° PCT/US2004/016964, publicado como o documento WO 2004/111964 A2, no qual o cálculo seleciona, a partir de um grupo de duas ou mais funções de modelo de intensidade acústica específica, uma ou uma combinação de duas ou mais funções de modelo de intensidade acústica específica, cuja seleção é controlada pela medida das características do sinal de áudio de entrada. A descrição da Intensidade acústica específica 104 da Figura 1, abaixo, descreve tal disposição.

[0085] De acordo com outros aspectos da presente invenção, o sinal de áudio não modificado ou (1) os parâmetros de modificação ou (2) a intensidade acústica alvo específica ou uma representação da intensidade acústica alvo específica (por exemplo, os fatores de escala utilizáveis no cálculo explícito ou implícito da intensidade acústica alvo específica) podem ser armazenados ou transmitidos para uso, por exemplo, em um dispositivo ou processo temporalmente e/ou espacialmente separados. Os parâmetros de modificação, a intensidade acústica alvo específica, ou a representação da intensidade acústica alvo específica podem ser determinados de uma maneira adequada, como, por exemplo, em um dentre os exemplos de disposição de alimentação direta, de alimentação, ou de alimentação dianteira e retroalimentação híbrida das Figuras 1 a 4, conforme descrito acima. Na prática, uma disposição de alimentação direta, como no exemplo da Figu-

ra 1, é a menos complicada e mais rápida, uma vez que evita os cálculos baseados no sinal de áudio modificado. Um exemplo de transmissão ou armazenamento do áudio não modificado e dos parâmetros de modificação é mostrado na Figura 5, enquanto um exemplo da transmissão ou armazenamento do áudio não modificado e da intensidade acústica alvo específica ou uma representação da intensidade acústica alvo específica é mostrado na Figura 6.

[0086] Uma disposição como a do exemplo da Figura 5 pode ser usada para separar temporal e/ou espacialmente a aplicação dos parâmetros de modificação ao sinal de áudio a partir da geração de tais parâmetros de modificação. Uma disposição tal como a da Figura 6 pode ser usada de modo a separar temporal e/ou espacialmente tanto a geração como a aplicação dos parâmetros de modificação da geração da intensidade acústica alvo específica ou sua representação. Ambos os tipos de disposições possibilitam uma disposição de reprodução ou recepção de baixo custo, evitando a complexidade de se gerar parâmetros de modificação ou de se gerar a intensidade acústica alvo específica. Embora um tipo de disposição da Figura 5 seja mais simples que o tipo de disposição da Figura 6, a disposição da Figura 6 tem a vantagem de a informação requerida para ser armazenada ou transmitida poder ser muito menor, particularmente quando uma representação da intensidade acústica específica, tal como um ou mais fatores de escala, é armazenada ou transmitida. Tal redução de armazenamento ou transmissão de informação pode ser particularmente útil em ambientes de áudio de baixa taxa de bits.

[0087] Por conseguinte, outros aspectos da presente invenção são a provisão de um dispositivo ou processo (1) que recebe ou reproduz, a partir de um armazenador ou dispositivo ou processo de transmissão, parâmetros de modificação M e aplica os mesmos a um sinal de áudio que é também recebido ou (2) que recebe ou reproduz, a partir

de um armazenador ou dispositivo ou processo de transmissão, uma intensidade acústica alvo específica ou representação de uma intensidade acústica alvo específica, gera parâmetros de modificação M por meio da aplicação da intensidade acústica alvo específica ou sua representação para o sinal de áudio que é também recebido (ou para uma medida do sinal de áudio tal como sua intensidade acústica específica, que pode ser obtida a partir do sinal de áudio), e aplica os parâmetros de modificação M para o sinal de áudio recebido. Estes dispositivos ou processos podem ser caracterizados como processos de decodificação ou decodificadores; enquanto os dispositivos ou processos requeridos para produzir as informações armazenadas ou transmitidas podem ser caracterizados como processos de codificação ou codificadores. Estes processos de codificação ou codificadores são as porções dos exemplos de disposição das Figuras 1 a 4 utilizáveis para se produzir as informações requeridas pelos respectivos processos de decodificação ou decodificadores. Estes processadores de decodificação ou decodificadores podem ser associados a ou operativos com virtualmente qualquer tipo de processo ou dispositivo que processe e/ou reproduza som.

[0088] Em um aspecto da presente invenção, como no exemplo da Figura 5, o sinal de áudio não modificado e os parâmetros de modificação M produzidos, por exemplo, por um processo de geração ou gerador de parâmetros de modificação, tais como o Gerar Parâmetros de modificação 4 da Figura 1, 4' da Figura 2, 4" da Figura 3 ou 4''' da Figura 4, podem ser aplicados a qualquer dispositivo ou função de armazenamento ou transmissão adequado ("Armazenar ou Transmitir") 16. No caso de se usar o exemplo de alimentação direta da Figura 1 como um processo de codificação ou codificador, o Modificar Sinal de áudio 2 não precisará gerar o áudio modificado e poderá ser omitido quando não houver nenhuma exigência no sentido de prover o áudio

modificado no local temporal ou espacial do codificador ou processo de codificação. O Armazenar ou Transmitir 16 pode incluir, por exemplo, qualquer dispositivo de armazenamento e reprodução magnético, óptico, ou de estado sólido ou qualquer dispositivo de transmissão e recepção de fio ou sem fio, a escolha dos mesmos não sendo uma questão crítica para a presente invenção. Os parâmetros de modificação reproduzidos ou recebidos podem então ser aplicados a um Modificar Sinal de áudio 2, do tipo empregado nos exemplos das Figuras 1 a 4, a fim de modificar o sinal de áudio reproduzido ou recebido de modo que sua intensidade acústica específica se aproxime da intensidade acústica alvo específica da ou inerente à disposição na qual os parâmetros de modificação são obtidos. Os parâmetros de modificação podem ser armazenados ou transmitidos de diversas maneiras. Por exemplo, eles podem ser armazenados ou transmitidos como metadados anexos ao sinal de áudio, os mesmos podem ser enviados em separados trajetos ou canais, podem ser codificados de maneira esteganográfica no áudio, podem ser multiplexados, etc. o uso dos parâmetros de modificação para modificar o sinal de áudio pode ser opcional e, se opcional, o seu uso pode ser selecionado, por exemplo, por um usuário. Por exemplo, os parâmetros de modificação, quando aplicados ao sinal de áudio, podem reduzir a faixa dinâmica do sinal de áudio. Empregar ou não tal redução de faixa dinâmica pode ser selecionado por um usuário.

[0089] Em um outro aspecto da presente invenção, como no exemplo da Figura 6, o sinal de áudio não modificado e a intensidade acústica alvo específica ou representação da intensidade acústica alvo específica pode ser aplicada a qualquer dispositivo ou função de armazenamento ou transmissão adequado ("Armazenar ou Transmitir") 16. No caso de se usar uma configuração de alimentação direta, tal como no exemplo da Figura 1, como um processo de codificação ou

um codificador, nem um processo ou dispositivo do tipo Calcular Parâmetros de modificação 10, nem um processo ou dispositivo do tipo Modificar Sinal de áudio 2 seria necessário e poderiam ser omitidos caso não haja nenhuma exigência no sentido prover parâmetros de modificação ou o áudio modificado no local temporal ou espacial do codificador ou processo de codificação. Como no caso da Figura 5, o Armazenar ou Transmitir 16 pode incluir, por exemplo, qualquer dispositivo de armazenamento e reprodução magnético, óptico, ou de estado sólido ou qualquer dispositivo de transmissão e recepção de fio ou sem fio adequado, a escolha dos mesmos não sendo uma questão crítica para a presente invenção. A intensidade acústica alvo específica reproduzida ou recebida ou representação da intensidade acústica alvo específica pode então ser aplicada, juntamente com o áudio não modificado, a um Calcular Parâmetros de modificação 10, do tipo empregado no exemplo da Figura 1, ou a um Calcular Parâmetros de modificação 10", do tipo empregado no exemplo da Figura 3, a fim de prover os parâmetros de modificação M que podem então ser aplicados a um Modificar Sinal de áudio 2, do tipo empregado nos exemplos das Figuras 1 a 4, a fim de modificar o sinal de áudio reproduzido ou recebido de modo que sua intensidade acústica específica se aproxime da intensidade acústica alvo específica da ou inerente à disposição na qual os parâmetros de modificação são obtidos. Embora a intensidade acústica alvo específica ou sua representação possa ser mais prontamente obtida em um processo de codificação ou codificador do tipo da Figura 1, a intensidade acústica alvo específica ou sua representação ou uma aproximação à intensidade acústica alvo específica ou sua representação pode ser obtida em um processo de codificação ou codificador das Figuras 2 a 4 (as aproximações são calculadas nos processos ou dispositivos 14 das Figuras 2 e 3 e no processo ou dispositivo 12 da Figura 4). A intensidade acústica alvo específica ou sua

representação pode ser armazenada ou transmitida de diversas maneiras. Por exemplo, a mesma pode ser armazenada ou transmitida como metadados anexados ao sinal de áudio, pode ser enviada em trajetos ou canais separados, pode ser codificada de maneira esteganográfica no áudio, pode ser multiplexada, etc. O uso dos parâmetros de modificação obtidos da intensidade acústica alvo específica armazenada ou transmitida ou representação para modificar o sinal de áudio pode ser opcional e, se opcional, o seu uso pode ser selecionado, por exemplo, por um usuário. Por exemplo, os parâmetros de modificação, quando aplicados ao sinal de áudio, podem reduzir a faixa dinâmica do sinal de áudio. Empregar ou não tal redução de faixa dinâmica pode ser selecionado por um usuário.

[0090] Ao se implementar a presente invenção como um sistema digital, uma configuração de alimentação direta é mais prática, e exemplos de tais configurações são, portanto, descritos abaixo em detalhe, ficando entendido que o âmbito de aplicação da presente invenção não se limita aos mesmos.

[0091] Por todo o presente documento, termos tais como "filtro" ou "banco de filtros" são usados de modo a incluir essencialmente qualquer forma de filtragem recursiva ou não recursiva, tal como os filtros IIR ou transformadores, e as informações "filtradas" são o resultado da aplicação de tais filtros. As modalidades descritas abaixo empregam bancos de filtro implementados por transformadores.

[0092] A Figura 7 ilustra detalhes de uma modalidade exemplar de um aspecto da presente invenção incorporada em uma disposição de alimentação direta. O áudio, primeiramente, passa por uma função ou dispositivo de banco de filtros de análise ("Banco de filtros de Análise") 100, que divide o sinal de áudio em uma pluralidade de bandas de frequência (assim, a Figura 5 mostra múltiplas saídas do Banco de filtros de Análise 100, cada saída representando uma banda de frequência,

cuja saída carrega as várias funções ou dispositivos até um banco de filtros de síntese, que soma as bandas em um sinal de banda larga combinado, conforme descrito em mais detalhe abaixo). A resposta do filtro associado a cada banda de frequência no Banco de filtros de Análise 100 é concebida de modo a simular a resposta em um local em particular da membrana basilar no ouvido interno. A saída de cada filtro no Banco de filtros de Análise 100, em seguida, passa para um filtro de transmissão ou função de filtro de transmissão ("Filtro de Transmissão") 101 que simula o efeito de filtragem da transmissão de áudio através do ouvido externo e médio. Se apenas a intensidade acústica do áudio tivesse de ser medida, o filtro de transmissão poderia ser aplicado antes do banco de filtros de análise, mas, uma vez que as saídas do banco de filtros de análise são usadas para sintetizar o áudio modificado, é vantajoso se aplicar o filtro de transmissão depois do banco de filtros. As saídas do Filtro de transmissão 101 em seguida passam para uma função ou dispositivo de excitação ("Excitação") 102, cujas saídas simulam a distribuição de energia ao longo da membrana basilar. Os valores da energia de excitação podem ser suavizados com o tempo por meio de uma função ou dispositivo de suavização ("Suavização") 103. As constantes do tempo da função de suavização são definidas de acordo com as exigências de uma aplicação desejada. Os sinais de excitação suavizados são em seguida convertidos em intensidade acústica específica na função ou dispositivo de intensidade acústica específica ("Intensidade acústica específica" (SL)) 104. A intensidade acústica específica é representada em unidades de sone por frequência unitária. O componente de intensidade acústica específica associado a cada banda passa para uma função ou dispositivo de modificação ou dispositivo de modificação de intensidade acústica específica ("Modificação de SL") 105. A modificação de SL 105 toma como sua entrada a intensidade acústica específica ori-

ginal e em seguida emite uma intensidade acústica específica "alvo" ou desejada, que, de acordo com um aspecto da presente invenção é de preferência uma função da intensidade acústica específica original (vide abaixo o parágrafo intitulado "Intensidade acústica alvo específica". A modificação de SL 105 pode operar de forma independente em cada banda, ou pode haver uma interdependência entre bandas (uma suavização de frequência, conforme sugerido pelas linhas cruzadas na Figura 7), dependendo do efeito desejado. Ao se tomar como suas entradas os componentes de banda de frequência de excitação suavizada da Excitação 102 e a intensidade acústica alvo específica da Modificação de SL 105, uma função ou dispositivo de solucionador de ganho ("Solucionador de Ganho") 106 determina o ganho que precisa ser aplicado a cada banda da saída do Banco de filtros de Análise 100 a fim de transformar a intensidade acústica específica medida na intensidade acústica alvo específica. O solucionador de ganho pode ser implementado de várias maneiras. Por exemplo, o Solucionador de ganho pode incluir um processo iterativo, tal como o apresentado no dito Pedido de Patente Internacional Nº PCT/US2004/016964, publicado como o documento WO 2004/111994 A2, ou de maneira alternativa, uma pesquisa de tabela. Embora os ganhos por banda gerados pelo Solucionador de Ganho 106 possam ser mais suavizados ao longo do tempo por meio da função ou dispositivo de suavização opcional ("Suavização") 107 a fim de minimizar os artefatos perceptivos, é preferido que a suavização temporal seja aplicada em qualquer lugar por todo o processo ou dispositivo, conforme descrito. Finalmente, os ganhos são aplicados às respectivas bandas do Banco de filtros de Análise 100 através de uma função de combinação ou combinador multiplicativa 108, e o áudio processado ou "modificado" é sintetizado a partir das bandas modificadas no ganho em uma função ou dispositivo de banco de filtros de síntese ("Banco de filtros de Síntese") 110. Além disso, as

saídas do banco de filtros de análise podem ser atrasadas por uma função ou dispositivo de retardo ("Retardo") 109 antes da aplicação dos ganhos a fim de compensar qualquer latência associada à computação de ganho. De maneira alternativa, ao invés de se calcular os ganhos para uso na aplicação de modificações de ganho em bandas de frequência, os Solucionadores de Ganho 106 podem calcular os coeficientes de filtro que controlam um filtro variante no tempo, tal como um filtro FIR de diversas derivas ou um filtro IIR multipolar. Para fins de simplicidade da exposição, aspectos da presente invenção são principalmente descritos empregando fatores de ganho aplicados às bandas de frequência, sendo entendido que os coeficientes de filtro e os filtros variantes no tempo podem também ser empregados em modalidades práticas.

[0093] Nas modalidades práticas, o processamento do áudio pode ser feito no domínio digital. Por conseguinte, o sinal de entrada de áudio é indicado pela sequência de tempo discreta $x[n]$ que é amostrada a partir da fonte de áudio em alguma frequência de amostragem f_s . Supõe-se que a sequência $x[n]$ seja apropriadamente graduada de modo que a força rms de $x[n]$ em decibéis dada por

$$RMS_{dB} = 10 \log_{10} \left(\frac{1}{L} \sum_{n=0}^{L-1} x^2[n] \right)$$

seja igual ao nível de pressão de som em dB no qual o áudio é ouvido por um ouvinte humano. Além disso, pressupõe-se que o sinal de áudio seja monofônico para fins de simplicidade de exposição.

[0094] O Banco de filtros de Análise 100, o Filtro de transmissão 101, a Excitação 102, a Intensidade acústica específica 104, a Modificação de Intensidade acústica específica 105, o Solucionador de Ganho 106, e o Banco de filtros de Síntese 110 podem ser descritos em mais detalhes como segue.

O Banco de filtros de Análise 100

[0095] O sinal de entrada de áudio é aplicado a um banco de filtros

de análise ou função de banco de filtros ("Banco de filtros de Análise") 100. Cada filtro do Banco de filtros de Análise 100 é concebido para simular a resposta de frequência em um local em particular ao longo da membrana basilar do ouvido interno. O banco de filtros 100 pode incluir um conjunto de filtros lineares cuja largura de banda e espaçamento são constantes na escala de frequência de Largura de banda Retangular Equivalente (ERB), como definido por Moore, Glasberg e Baer (B. C. J. Moore, B. Glasberg, T. Baer, "A Model for the Prediction of Thresholds, Loudness, and Partial Loudness", supra).

[0096] Embora a escala de frequência ERB combine mais proximamente com a percepção humana e mostre um desempenho melhor na produção de medições de intensidade acústica objetiva que correspondem aos resultados de intensidade acústica subjetiva, a escala de frequência de Bark pode ser empregada com menor desempenho.

[0097] Para uma frequência central f em hertz, a largura de uma banda ERB em hertz pode ser aproximada como:

$$\text{ERB}(f) = 24.7 (4.37 f / 1000 + 1) \quad (1)$$

[0098] A partir desta relação uma escala de frequência deformada é definida de tal modo que em qualquer ponto ao longo da escala deformada, a ERB correspondente em unidades da escala deformada seja igual a um. A função de conversão de frequência linear em hertz nesta escala de frequência ERB é obtida por meio da integração da recíproca da Equação 1:

$$\text{HzToERB}(f) = \int \frac{1}{24.7(4.37 f / 1000 + 1)} df = 21.4 \log_{10}(4.37 f / 1000 + 1) \quad (2a)$$

[0099] É igualmente útil expressar a transformação da escala ERB novamente na escala de frequência linear por meio da solução da Equação 2a para f :

$$\text{ERBToHz}(e) = f = \frac{1000}{4.37} 10^{(e/21.4-1)}, \quad (2b)$$

na qual e é em unidades da escala ERB. A Figura 9 mostra a relação

entre a escala ERB e a frequência em hertz.

[00100] O Banco de filtros de Análise 100 pode incluir os filtros auditivos B, referidos como bandas, nas frequências centrais $f_c [1] \dots f_c [B]$ espaçadas uniformemente ao longo da escala ERB. Mais especificamente,

$$f_c[1] = f_{\min} \quad (3a)$$

$$f_c[b] = f_c[b-1] + \text{ERBToHz}(\text{HzToERB}(f_c[b-1]) + \Delta) \quad b = 2 \dots B \quad (3b)$$

$$f_c[B] < f_{\max} \quad (3c)$$

em que Δ é o espaçamento ERB desejado do Banco de filtros de Análise 100, e nas quais $f_{\min}$ e $f_{\max}$ são as frequências centrais mínima e máxima desejadas, respectivamente. Pode-se escolher $\Delta = 1$, levando-se em consideração a faixa de frequência sobre a qual o ouvido humano é sensível, e pode-se definir $f_{\min} = 50$ Hz e $f_{\max} = 20.000$ Hz. Com tais parâmetros, por exemplo, a aplicação das Equações 3a a c produz $B = 40$ filtros auditivos.

[00101] A resposta de frequência de magnitude de cada filtro auditivo pode ser caracterizado por uma função exponencial arredondada, conforme sugerido por Moore e Glasberg. Em termos específicos, a resposta de magnitude de um filtro com uma frequência central $f_c [b]$ pode ser computada como:

$$H_b(f) = (1 + pg)e^{-pg} \quad (4a)$$

$$g = \left| \frac{f - f_c[b]}{f_c[b]} \right|, \quad (4b)$$

$$p = \frac{4f_c[b]}{\text{ERB}(f_c[b])} \quad (4c)$$

[00102] Nas quais as respostas de magnitude de tais filtros auditivos B que se aproxima da formação de banda crítica na escala ERB são mostradas na Figura 10.

[00103] As operações de filtragem do Banco de filtros de Análise 100 podem ser adequadamente aproximadas usando uma Transformação Discreta de Fourier de comprimento finito, normalmente referi-

da como a Transformação Discreta de Fourier de Tempo Curto (STDFT), uma vez que se supõe que uma implementação executada nos filtros na taxa de amostragem do sinal de áudio, referida como uma implementação taxa plena ("full-rate"), provenha uma resolução mais temporal que o necessário para precisas medições de intensidade acústica. Ao se usar a STDFT ao invés de uma implementação taxa plena ("full-rate"), um aumento da eficiência e uma redução da complexidade computacional podem ser obtidas.

[00104] A transformação STDFT do sinal de áudio de entrada $x[n]$ é definida como:

$$X[k, t] = \sum_{n=0}^{N-1} w[n] x[n + tT] e^{-j \frac{2\pi k n}{N}}, \quad (5a)$$

na qual k é o índice de frequência, t é o índice de bloco de tempo, N é o tamanho da DFT, T é o tamanho do segmento hop, e $w[n]$ é uma janela de tamanho N normalizada de modo que

$$\sum_{n=0}^{N-1} w^2[n] = 1 \quad (5b)$$

[00105] Observa-se que a variável t na Equação 5a é um índice discreto representando o bloco de tempo da STDFT em oposição a uma medida de tempo em segundos. Cada incremento em t representa um segmento hop de amostras T ao longo do sinal $x[n]$. As referências subsequentes ao índice t pressupõem esta definição. Embora diferentes definições de parâmetros e formas de janelas possam ser usadas dependendo dos detalhes da implementação, para $f_s = 44100$ Hz, escolher $N = 2048$, $T = 1024$, e ter $w[n]$ como uma janela de Hanning oferece um equilíbrio de tempo e resolução de frequência adequados. A STDFT descrita acima pode ser mais eficiente utilizando a Transformação Rápida de Fourier (FFT).

[00106] Ao invés da STDFT, a Transformação Discreta de Co-seno Modificada (MDCT) pode ser utilizada para implementar o banco de filtros de análise. A MDCT é uma transformação normalmente usada

em codificadores de áudio perceptivos, tais como os Dolby AC-3. Quando o sistema apresentado é implementado com tal áudio codificado perceptivamente, a medição e a modificação da intensidade acústica apresentadas poderão mais eficientemente implementadas por meio do processamento dos coeficientes da MDCT existentes do áudio codificado, desta forma eliminando a necessidade de se fazer a transformação do banco de filtros de análise. A MDCT do sinal de áudio de entrada $x[n]$ é calculada por:

$$X[k, t] = \sum_{n=0}^{N-1} w[n] x[n + tT] \cos((2\pi / N)(k + 1/2)(n + n_0)),$$

na qual $n_0 = \frac{(N/2) + 1}{2}$ (6)

[00107] De modo geral, o tamanho do segmento hop T é exatamente a metade do comprimento N da transformação de modo a ser possível uma perfeita reconstrução do sinal $x[n]$.

Filtro de Transmissão 101

[00108] As saídas do Banco de filtros de Análise 100 são aplicadas a um filtro de transmissão ou a uma função de filtro de transmissão ("Filtro de transmissão") 101 que filtra cada banda do banco de filtros de acordo com a transmissão do áudio através do ouvido externo e intermediário. A Figura 8 ilustra uma resposta de frequência de magnitude adequada do filtro de transmissão, $P(f)$, através da faixa de frequência audível. A resposta é uma unidade abaixo de 1 kHz, e acima de 1 kHz, seguindo o inverso do limite da audição conforme especificado no padrão ISO226, com o limite normalizado em unidade igual a 1 kHz.

Excitação 102

[00109] A fim de computar a intensidade acústica do sinal de áudio de entrada, uma medição da energia passageira dos sinais de áudio em cada filtro do Banco de filtros de Análise 100 após a aplicação do Filtro de transmissão 101 torna-se necessária. Esta medição variante

no tempo e na frequência é referida como excitação. A saída de energia passageira de cada filtro no Banco de filtros de Análise 100 pode ser aproximada na Função de Excitação 102 através da multiplicação das respostas dos filtros no domínio da frequência com o espectro de força do sinal de entrada:

$$E[b, t] = \frac{1}{N} \sum_{k=0}^{N-1} |H_b[k]|^2 |P[k]|^2 |X[k, t]|^2, \quad (7)$$

na qual b é o número de banda, t é o número de bloco, e $H_b[k]$ e $P[k]$ são as respostas de frequência do filtro auditivo e do filtro de transmissão, respectivamente, amostrados em uma frequência correspondente ao índice k binário da transformação STDFT ou MDCT. Deve-se notar que as formas para a resposta de magnitude dos filtros auditivos diferentes das especificadas nas Equações 4a a 4c podem ser usadas na Equação 7 para se obter resultados similares. Por exemplo, o dito Pedido Internacional N° PCT/US2004/016964, publicado como o documento WO 2004/111994 A2, descreve duas alternativas: um filtro auditivo caracterizado por uma função de transferência de 12ª ordem, e uma aproximação passa banda de "parede de tijolo" de baixo custo.

[00110] Em suma, a saída da Função de Excitação 102 é uma representação de domínio de frequência da energia E nas respectivas bandas ERB por período de tempo t .

Intervalo de Tempo ("Suavização") 103

[00111] Para certas aplicações da presente invenção, conforme abaixo descrito, pode ser desejável suavizar a excitação $E[b, t]$ antes de sua transformação em intensidade acústica específica. Por exemplo, a suavização pode ser feita recursivamente na função Suavização 103 de acordo com a equação:

$$\bar{E}[b, t] = \lambda_b \bar{E}[b, t] + (1 - \lambda_b) E[b, t], \quad (8)$$

na qual as constantes de tempo λ_b em cada banda b são selecionadas de acordo com a aplicação desejada. Na maioria dos casos, as cons-

tantes de tempo podem ser vantajosamente proporcionais ao tempo de integração da percepção humana da intensidade acústica dentro da banda b . Watson e Gengel realizaram experimentos demonstrando que este tempo de integração está dentro da faixa de 150 a 175 ms a baixas frequências (125 a 200 Hz) e de 40 a 60 ms a altas frequências (Charles S. Watson e Roy W. Gengel, "Signal Duration and Signal Frequency in Relation to Auditory Sensitivity" *Journal of the Acoustical Society of America*, Vol. 46, N. 4 (Parte 2), 1969, pp. 989-997).

Intensidade acústica específica 104

[00112] No conversor ou função de conversão da intensidade acústica específica ("Intensidade acústica específica") 104, cada banda de frequência é convertida em um valor componente da intensidade acústica específica, medida em sone por ERB.

[00113] Inicialmente, na computação da intensidade acústica específica, o nível de excitação em cada banda de $E[b, t]$ pode ser transformado em um nível de excitação equivalente a 1 kHz, conforme especificado pelos contornos iguais de intensidade acústica do ISO 226 (Figura 11) normalizados pelo filtro de transmissão P9z) (Figura 12):

$$\bar{E}_{1kHz}[b, t] = T_{1kHz}(E[b, t], f_c[b]), \quad (9)$$

na qual $T_{1kHz}(E, f)$ é uma função que gera o nível a 1 kHz, o que é igualmente alto para o nível E na frequência f . Na prática, $T_{1kHz}(E, f)$ é implementado como uma interpolação de uma tabela de pesquisa dos contornos iguais da intensidade acústica, normalizados pelo filtro de transmissão. A transformação a níveis equivalentes a 1 kHz simplifica o seguinte cálculo de intensidade acústica específica.

[00114] Em seguida, a intensidade acústica específica em cada banda pode ser computada como:

$$N[b, t] = \alpha[b, t]N_{NB}[b, t] + (1 - \alpha[b, t])N_{WB}[b, t], \quad (10)$$

na qual $N_{NB}[b, t]$ e $N_{WB}[b, t]$ são valores de intensidade acústica específica baseados em um modelo de sinal de banda estreita e banda lar-

ga, respectivamente. O valor $E[b, t]$ é um fator de interpolação entre 0 e 1 computado a partir do sinal de áudio. O dito Pedido PCT/US2004/016964, publicado como o documento WO 2004/111994 A2, descreve uma técnica para o cálculo de $E[b, t]$ a partir do achatamento espectral da excitação. O documento descreve ainda os modelos de sinal de "banda estreita" e de "banda larga" em mais detalhes.

[00115] Os valores de intensidade acústica de banda estreita e banda larga $N_{NB}[b, t]$ e $N_{WB}[b, t]$ podem ser estimados a partir da excitação transformada usando as funções exponenciais:

$$N_{NB}[b, t] = \begin{cases} G_{NB} \left(\left(\frac{\bar{E}_{1kHz}[b, t]}{TQ_{1kHz}} \right)^{\alpha_{NB}} - 1 \right), & \bar{E}_{1kHz}[b, t] > TQ_{1kHz} \\ 0, & otherwise \end{cases} \quad (11a)$$

$$N_{WB}[b, t] = \begin{cases} G_{WB} \left(\left(\frac{\bar{E}_{1kHz}[b, t]}{TQ_{1kHz}} \right)^{\alpha_{WB}} - 1 \right), & \bar{E}_{1kHz}[b, t] > TQ_{1kHz} \\ 0, & otherwise \end{cases}, \quad (11b)$$

nas quais TQ_{1kHz} é o nível de excitação no limite em silêncio para um tom de 1 kHz. A partir dos contornos de intensidade acústica iguais (Figuras 11 e 12), TQ_{1kHz} é igual a 4,2 dB. Nota-se que ambas as funções de intensidade acústica específica são iguais a zero quando a excitação é igual ao limite em silêncio. Para excitações maiores que o limite em silêncio, ambas as funções crescem monotonamente com uma lei de força de acordo com a lei de Stevens de sensação de intensidade. O expoente para a função de banda estreita é maior que o da função de banda larga, fazendo a função de banda estreita aumentar mais rapidamente que a função de banda larga. A seleção específica dos expoentes β e os ganhos G para os casos de banda estreita e banda larga são escolhidos de modo a corresponderem aos dados experimentais sobre o crescimento da intensidade acústica para tons e ruído.

[00116] Moore e Glasberg sugerem que a intensidade acústica específica deve ser igual a algum valor pequeno ao invés de zero quan-

do a excitação está no limite da audição. A intensidade acústica específica deve então diminuir monotonamente a zero à medida que a excitação diminui para zero. A justificativa é que o limite da audição é um limite probabilístico (o ponto no qual um tom é detectado 50 % das vezes), e que vários tons, cada qual no limite, apresentados juntos, podem somar um som mais audível que qualquer um dos tons individuais. Na aplicação apresentada, aumentar as funções de intensidade acústica específica com esta propriedade tem o benefício adicional de fazer com que o solucionador de ganho, apresentado abaixo, se comporte mais apropriadamente quando a excitação está perto do limite. Quando a intensidade acústica específica é definida como zero, quando a excitação está no ou abaixo do limite, uma solução única para o solucionador de ganho não existirá para as excitações no ou abaixo do limite. Quando, por outro lado, a intensidade acústica específica é definida de modo a aumentar monotonamente em todos os valores de excitação maior que ou igual a zero, conforme sugerido por Moore e Glasberg, uma solução única existirá. Uma graduação de intensidade acústica maior que a unidade sempre resultará em um ganho maior que a unidade ou vice-versa. As funções de intensidade acústica específica nas Equações 11a e 11b podem ser alteradas de modo a apresentar uma propriedade desejada de acordo com:

$$N_{NB}[b, t] = \begin{cases} G_{NB} \left(\left(\frac{\bar{E}_{1kHz}[b, t]}{TQ_{1kHz}} \right)^{\eta_{NB}} - 1 \right), & \bar{E}_{1kHz}[b, t] > \lambda TQ_{1kHz} \\ \exp \left[K_{NB} \left(-\log \left(\frac{\bar{E}_{1kHz}[b, t]}{TQ_{1kHz}} \right) + C_{NB} \right)^{\eta_{NB}} \right], & otherwise \end{cases} \quad (11c)$$

$$N_{NB}[m, t] = \begin{cases} G_{NB} \left(\left(\frac{\bar{E}_{1kHz}[b, t]}{TQ_{1kHz}} \right)^{\eta_{NB}} - 1 \right), & \bar{E}_{1kHz}[b, t] > \lambda TQ_{1kHz} \\ \exp \left[K_{NB} \left(-\log \left(\frac{\bar{E}_{1kHz}[b, t]}{TQ_{1kHz}} \right) + C_{NB} \right)^{\eta_{NB}} \right], & otherwise \end{cases} \quad (11d)$$

nas quais a constante λ é maior que um, o expoente η é menor que um, e as constantes K e C são escolhidas de tal modo que a função de

intensidade acústica específica e sua primeira derivada sejam contínuas no ponto $E_{\text{kHz}}[b, t] = \lambda TQ_{\text{kHz}}$.

[00117] A partir da intensidade acústica específica, a intensidade acústica geral ou "total" $L[t]$ é dada pela soma da intensidade acústica específica ao longo de todas as bandas b :

$$L[t] = \sum_b N[b, t] \quad (12)$$

Modificação da Intensidade acústica específica 105

[00118] Na função de modificação de intensidade acústica específica ("Modificação de Intensidade acústica específica") 105, a intensidade acústica alvo específica, referida como $N[b, t]$, pode ser calculada a partir da intensidade acústica específica do nível SL 104 (Figura 7) de várias maneiras dependendo da aplicação desejada do dispositivo ou processo geral. Conforme descrito em mais detalhes abaixo, uma intensidade acústica alvo específica pode ser calculada usando um fator de escala α , por exemplo, no caso de um controle de volume. Vide Equação 16 abaixo e sua descrição associada. No caso do controle de ganho automático (AGC) e do controle de faixa dinâmica (DRC), uma intensidade acústica alvo específica pode ser calculada usando uma razão de intensidade acústica de saída desejada para a intensidade acústica de entrada. Vide as Equações 17 e 18 abaixo e suas descrições associadas. No caso de uma equalização dinâmica, uma intensidade acústica alvo específica pode ser calculada usando uma relação apresentada na Equação 23 e sua descrição associada.

Solucionador de Ganho 106

[00119] Neste exemplo, para cada banda b e cada intervalo de tempo t , o Solucionador de ganho 106 toma como suas entradas a excitação suavizada $E[b, t]$ e a intensidade acústica alvo específica $N[b, t]$ e gera os ganhos $G[b, t]$ usados em seguida para a modificação do áudio. A função $\psi\{\cdot\}$ representa a transformação não linear a partir da excitação para a intensidade acústica específica de modo que:

$$N[b, t] = \Psi\{\bar{E}[b, t]\}, \quad (13)$$

e o Solucionador de ganho encontra $G[b, t]$ de modo que:

$$\hat{N}[b, t] = \Psi\{G^2[b, t]\bar{E}[b, t]\}. \quad (14a)$$

[00120] Os Solucionadores de ganho 106 determinam os ganhos variantes na frequência e no tempo, que, quando aplicados à excitação original, resultam em uma intensidade acústica específica que, em termos ideais, é igual à intensidade acústica alvo específica desejada. Na prática, os Solucionadores de ganho determinam os ganhos variantes na frequência e no tempo que, quando aplicados à versão de domínio da frequência do sinal de áudio, resultam na modificação do sinal de áudio a fim de reduzir a diferença entre a sua intensidade acústica específica e a intensidade acústica alvo específica. Idealmente, a modificação é tal que o sinal de áudio modificado tenha uma intensidade acústica específica que é uma aproximação estrita da intensidade acústica alvo específica. A solução para a Equação 14a pode ser implementada de várias maneiras. Por exemplo, quando há uma expressão matemática de sentença fechada para o inverso da intensidade acústica específica, representada por $\Psi^{-1}\{.\}$, os ganhos poderão ser computados diretamente pela equação reajustada 14a:

$$G[b, t] = \sqrt{\frac{\Psi^{-1}\{\hat{N}[b, t]\}}{\bar{E}[b, t]}} \quad (14b)$$

[00121] De maneira alternativa, quando uma solução de sentença fechada para $\Psi^{-1}\{.\}$ não existe, uma abordagem iterativa poderá ser empregada na qual cada equação de iteração 14a é avaliada usando uma estimativa corrente dos ganhos. A intensidade acústica específica resultante é comparada ao alvo desejado e os ganhos são atualizados com base no erro. Quando os ganhos são atualizados apropriadamente, eles convergem para a solução desejada. Um outro método envolve a pré-computação da função $\Psi\{.\}$ para uma faixa de valores de excitação em cada banda a fim de criar uma tabela de pesquisa. A partir

desta tabela de pesquisa, obtém-se uma aproximação da função inversa $\psi^{-1}\{.\}$ e os ganhos podem ser computados a partir da equação 14b. conforme acima mencionado, a intensidade acústica alvo específica pode ser representada por uma graduação da intensidade acústica específica:

$$\hat{N}[b, t] = \Xi[b, t]N[b, t] \quad (14c)$$

[00122] Substituindo a equação 13 pela 14c e em seguida a 14c pela 14b, produz-se uma expressão alternativa para os ganhos:

$$G[b, t] = \sqrt{\frac{\Psi^{-1}(\Xi[b, t]\Psi(\bar{E}[b, t]))}{\bar{E}[b, t]}} \quad (14d)$$

[00123] Observa-se que os ganhos podem ser expressos basicamente como uma função da excitação $\bar{E}[b, t]$ e da graduação da intensidade acústica específica $\Xi[b, t]$. portanto, os ganhos podem ser computados através da avaliação da equação 14d ou de uma tabela de pesquisa equivalente sem nunca explicitamente computar a intensidade acústica específica ou a intensidade acústica alvo específica como valores intermediários. No entanto, estes valores são implicitamente computados através do uso da equação 14d. Outros métodos equivalentes para a computação dos parâmetros de modificação através de uma computação explícita ou implícita da intensidade acústica específica e da intensidade acústica alvo específica podem ser considerados, e a presente invenção pretende abranger todos estes métodos.

Banco de filtros de Síntese 110

[00124] Conforme acima descrito, o Banco de filtros de Análise 100 pode ser implementado com eficiência através do uso da Transformação Discreta de Fourier de curto tempo (STDFT) ou da Transformação Discreta de Cosseno Modificada, e a STDFT ou a MDCT podem ser usadas de maneira similar a fim de implementar o Banco de filtros de Síntese 110. Em termos específicos, ao se deixar $X[k, t]$ representar a

STDFT ou a MDCT do áudio de entrada, conforme acima descrito, a STDFT ou a MDCT do áudio (modificado) processado no Banco de filtros de Síntese 110 podem ser calculadas como:

$$\hat{X}[k, t] = \sum_b G[b, t] S_b[k] X[k, t - d], \quad (15)$$

na qual $S_b[k]$ é a resposta do filtro de síntese associado à banda b , e d é o retardo associado ao bloco de retardos 109 da Figura 7. A forma dos filtros de síntese $S_b[k]$ pode ser igual a dos filtros utilizados no banco de filtros de análise, $H_b[k]$, ou pode ser modificada de modo a prover uma perfeita reconstrução na ausência de qualquer modificação de ganho (isto é, quando $G[b, t] = 1$). O áudio processado final pode então ser gerado através da transformação inversa de Fourier ou da transformação de cosseno modificada de $X[k, t]$ ou da síntese de adição sobreposta, conforme familiar a uma pessoa versada na técnica.

Intensidade acústica alvo específica

[00125] O comportamento das disposições que abrangem os aspectos da presente invenção, como, por exemplo, os exemplos das Figuras 1 a 7, é determinado principalmente pela maneira na qual a intensidade acústica alvo específica $N[b, t]$ é calculada. Embora a presente invenção não se limite a nenhuma função ou função inversa em particular para calcular a intensidade acústica alvo específica, diversas funções e aplicações adequadas para as mesmas serão descritas a seguir.

Função Invariante no tempo e Invariante na frequência Adequada para o Controle de Volume

[00126] Um controle de volume padrão ajusta a intensidade acústica de um sinal de áudio por meio da aplicação de um ganho de banda larga ao áudio. De modo geral, o ganho é acoplado a um botão ou corrediça que é ajustado por um usuário até que a intensidade acústica do áudio esteja em um nível desejado. Um aspecto da presente invenção permite uma maneira mais consistente em termos psicoacústicos

de se implementar tal controle. De acordo com este aspecto da presente invenção, ao invés de se ter um ganho de banda larga acoplado ao controle de volume que resulta em uma mudança do ganho pela mesma quantidade ao longo de todas as bandas de frequência, o que pode provocar uma mudança no espectro percebido, um fator de gradação de intensidade acústica específica é associado ao ajuste de controle de volume de modo que o ganho em cada uma dentre as múltiplas bandas de frequência mude em quantidade, levando em consideração o modelo de audição humana, e de modo que idealmente não haja uma mudança no espectro percebido. No contexto deste aspecto da presente invenção e em uma aplicação exemplar da mesma, "constante" ou "invariante no tempo" pretende permitir mudanças na definição de um fator de escala de controle de volume de tempos em tempos por um usuário, por exemplo. Esta invariância no tempo é às vezes referida como "quase invariante no tempo", "quase estacionária", "um pouco invariante no tempo", "um pouco estacionária", "invariante no tempo em escala", e "estacionária em escala". Dado tal fator de escala, α , a intensidade acústica alvo específica pode ser calculada como a intensidade acústica específica medida multiplicada por α :

$$\hat{N}[b, t] = \alpha N[b, t]. \quad (16)$$

[00127] Uma vez que a intensidade acústica total $L[t]$ é a soma da intensidade acústica específica $N[b, t]$ em todas as bandas b , a modificação acima também gradua a intensidade acústica total por um fator de α , mas de uma maneira a preservar o mesmo espectro percebido em um momento em particular para as mudanças no ajuste de controle de volume. Ou seja, em qualquer momento em particular, uma mudança no ajuste de controle de volume resulta em uma mudança na intensidade acústica percebida, porém nenhuma mudança no espectro percebido do áudio modificado versus o espectro percebido do áudio não modificado. A Figura 13a os ganhos multibanda resultantes $G[b, t]$

em todas as bandas b em um momento particular t quando $\alpha = 0,25$ para um sinal de áudio que consiste em uma fala feminina. Para fins de comparação, o ganho de banda larga requerido para graduar a intensidade acústica total original em 0,25 (linha horizontal), como em um controle de volume padrão, é também registrado. O ganho multibanda $G[b, t]$ aumenta nas bandas de baixa e alta frequências em comparação às bandas de frequência intermediárias. Isto é consistente com os contornos iguais de intensidade acústica, indicando que o ouvido humano é menos sensível a baixa e alta frequências.

[00128] A Figura 13b ilustra a intensidade acústica específica para o sinal de áudio original, o sinal modificado no ganho de banda larga conforme modificado de acordo com um controle de volume da técnica anterior, e o sinal modificado no ganho multibanda conforme modificado de acordo com este aspecto da presente invenção. A intensidade acústica específica do sinal modificado de ganho multibanda é o graduado original em 0,25. A intensidade acústica específica do sinal modificado em banda larga muda em forma espectral com relação à do sinal não modificado original. Neste caso, a intensidade acústica específica, em um sentido relativo, perde intensidade acústica em ambas baixa e alta frequências. Isto é percebido como uma surdez do áudio, uma vez que seu volume está desligado, um problema que não ocorre com o sinal modificado multibanda, cuja intensidade acústica é controlada pelos ganhos obtidos no domínio de intensidade acústica perceptiva.

[00129] Juntamente com a distorção do equipamento espectral percebido associado ao controle de volume tradicional, existe um segundo problema. Uma propriedade da percepção da intensidade acústica, que se reflete no modelo de intensidade acústica mostrado nas Equações 11a a 11d, é a intensidade acústica de um sinal em qualquer frequência que diminui mais rapidamente à medida que o nível de sinal

se aproxima do limite da audição. Como resultado, a atenuação elétrica requerida de modo a comunicar a mesma atenuação de intensidade acústica a um sinal mais suave será menor que a requerida para um sinal mais alto. Um controle de volume tradicional comunica atenuação, independentemente do nível de sinal, e, portanto, sinais suaves se tornam "suaves demais" com relação aos sinais mais altos quando o volume é desligado. Em muitos casos, isto resulta na perda do detalhe do áudio. Considere-se o registro de uma castanholha em um ambiente reverberante. Em tal registro, a batida principal da castanholha é bem alta em comparação aos ecos reverberantes, mas são os ecos reverberantes que dão o tamanho do ambiente. Quando o volume é desligado com um controle de volume tradicional, os ecos reverberantes ficam mais suaves com relação à batida principal e eventualmente desaparecem abaixo do limite da audição, deixando uma castanholha de som "seco". O controle de volume com base na intensidade acústica impede o desaparecimento das partes mais suaves das gravações ao reforçar a porção reverberante mais suave da gravação com relação à batida principal mais alta de modo que a intensidade acústica relativa entre estas seções permaneça constante. Para chegar a este efeito, os ganhos multibanda $G[b, t]$ devem variar ao longo do tempo a uma velocidade proporcional à resolução temporal humana da percepção da intensidade acústica. Uma vez que os ganhos multibanda $G[b, t]$ são computados como uma função da excitação suavizada $E[b, t]$, a seleção das constantes de tempo λ na Equação 8 determina quão rapidamente os ganhos poderão variar no tempo em cada banda b . conforme acima mencionado, estas constantes de tempo podem ser selecionadas de modo a serem proporcionais ao tempo de integração da percepção da intensidade acústica humana dentro da banda b e, deste modo, produzem a variação apropriada de $g[b, t]$ ao longo do tempo. Deve-se notar que se as constantes de tempo forem escolhidas de

maneira imprópria (ou rápido demais ou lento demais), artefatos perceptivamente objetáveis poderão ser introduzidos no áudio processado.

Função Invariante no tempo e Variante na frequência Adequada para Equalização Fixa

[00130] Em algumas aplicações, pode-se querer aplicar uma equalização perceptiva fixa ao áudio, em cujo caso a intensidade acústica alvo específica pode ser computada por meio da aplicação de um fator de escala invariante no tempo porém variante na frequência $\Theta[b]$ como na relação

$$\hat{N}[b, t] = \Xi[b, t] N[b, t],$$

na qual $(N[b, t])$ é a intensidade acústica alvo específica, $(N[b, t])$ é a intensidade acústica específica do sinal de áudio, b é uma medida de frequência, e t é uma medida de tempo. Neste caso, a graduação pode variar de banda para banda. Tal aplicação pode ser útil para enfatizar, por exemplo, a porção do espectro dominado pelas frequências de fala a fim de reforçar a inteligibilidade.

Função Invariante na frequência e Variante no tempo Adequada para Ganho Automático e Controle de Faixa Dinâmica

[00131] As técnicas do ganho automático e do controle de faixa dinâmica (ACG e DRC) são bem conhecidas no campo do processamento automático. Em um sentido abstrato, ambas as técnicas medem o nível de um sinal de áudio de alguma maneira e em seguida modificam em ganho o sinal em uma quantidade que vem a ser uma função do nível medido. No caso do AGC, o sinal é modificado em ganho de modo que o seu nível medido fique mais próximo de um nível de referência selecionado por um usuário. Com o controle DRC, o sinal é modificado em ganho de modo que a faixa do nível medido do sinal se transforme em uma faixa desejada. Por exemplo, pode-se querer fazer porções silenciosas de um áudio mais alto e as porções altas mais si-

lenciosas. Este sistema é descrito por Robinson e Gundry (Charles Robinson e Kenneth Gundry, "Dynamic Range Control via Metadata", 107th Convention of the AES, Preprint 5028, 24 a 27 de setembro de 1999, Nova Iorque). Implementações tradicionais do ganho AGC e do controle DRC de modo geral utilizam uma medição simples de nível de sinal de áudio, tal como pico suavizado ou amplitude de raiz quadrada mínima (rms), a fim de realizar a modificação de ganho. Estas medições simples se correlacionam em algum grau à intensidade acústica percebida do áudio, porém aspectos da presente invenção permitem um ganho AGC e um controle DRC mais relevantes perceptivamente ao fazer as modificações de ganho com uma medida de intensidade acústica com base em um modelo psicoacústico. Ainda, muitos sistemas de ganho AGC e de controle DRC tradicionais aplicam a modificação de ganho com um ganho de banda larga, incorrendo, assim, nas distorções timbrais (espectrais) acima mencionadas no áudio processado. Os aspectos da presente invenção, por outro lado, utilizam um ganho multibanda para formar a intensidade acústica específica de uma maneira a reduzir ou minimizar tais distorções.

[00132] Ambas as aplicações de AGC e DRC empregam aspectos da presente invenção caracterizados por uma função que transforma ou mapeia uma intensidade acústica de banda larga de entrada $L_i[t]$ em uma intensidade acústica de banda larga de saída desejada $L_o[t]$, na qual a intensidade acústica é medida em unidades de intensidade acústica perceptiva, tal como sone. A intensidade acústica de banda larga de entrada $L_i[t]$ é uma função da intensidade acústica específica do sinal de áudio de energia $N[b, t]$. Embora ela possa ser igual à intensidade acústica total do sinal de áudio de entrada, a mesma poderá ser uma versão suavizada temporalmente da intensidade acústica total do sinal de áudio.

[00133] As Figuras 14a e 14b ilustram exemplos de funções de ma-

peamento típicas de um AGC e de um DRC, respectivamente. Dado tal mapeamento no qual $L_o[t]$ é uma função de $L_i[t]$, a intensidade acústica alvo específica pode ser calculada como:

$$\hat{N}[b, t] = \frac{L_o[t]}{L_i[t]} N[b, t]. \quad (17)$$

[00134] A intensidade acústica específica original do sinal de áudio $N[b, t]$ é simplesmente graduada pela razão da intensidade acústica de banda larga de saída desejada para a intensidade acústica de banda larga de entrada de modo a produzir uma intensidade acústica específica de saída $N[b, t]$. Para um sistema de ganho AGC, a intensidade acústica de banda larga de entrada $L_i[t]$ deve de modo geral ser uma medida da intensidade acústica total de longo prazo do áudio. Isto pode ser obtido por meio da suavização da intensidade acústica total $L[t]$ ao longo do tempo a fim de gerar $L_i[t]$.

[00135] Em comparação a um ganho AGC, um sistema de controle DRC reage às mudanças de prazo menor em uma intensidade acústica do sinal, e, portanto, $L_i[t]$ poderá simplesmente ficar igual a $L[t]$. Como um resultado, a graduação da intensidade acústica específica, dada por $L_o[t] / L_i[t]$, pode flutuar rapidamente, resultando em artefatos indesejados no áudio processado. Um artefato típico é a modulação audível de uma porção do espectro de frequência por alguma parte não relacionada do espectro. Por exemplo, a seleção de uma música clássica deve conter altas frequências dominadas por uma nota de corda sustentada, enquanto as baixas frequências contêm um tímpano batendo alto. Sempre que o tímpano é batido, a intensidade acústica geral $L_i[t]$ aumenta, e o sistema DRC aplica uma atenuação a toda a intensidade acústica específica. As cordas são então ouvidas, subindo e descendo em intensidade acústica com o tímpano. Este bombeamento cruzado no espectro é um problema com os sistemas de DRC de banda larga tradicionais também, e uma solução típica envolve a aplicação de um DRC independente para diferentes bandas de fre-

quência diferentes. O sistema aqui apresentado é inerentemente multibanda devido ao banco de filtros e ao cálculo de uma intensidade acústica específica que emprega um modelo de intensidade acústica perceptiva, e, portanto, a modificação de um sistema de DRC a fim de operar de uma forma em multibanda de acordo com os aspectos da presente invenção é relativamente direta e é descrita a seguir.

Função Variante na frequência e Variante no tempo Adequada para Controle de Faixa Dinâmica

[00136] O sistema DRC pode ser expandido de modo a operar de uma maneira multibanda ou variante na frequência ao permitir que a intensidade acústica de entrada e saída varie independentemente com a banda b . estes valores de intensidade acústica multibanda são referidos como $L_i [b, t]$ e $L_o [b, t]$, e a intensidade acústica alvo específica pode em seguida ser dada por:

$$\hat{N}[b, t] = \frac{L_o[b, t]}{L_i[b, t]} N[b, t], \quad (18)$$

em que $L_o [b, t]$ é calculado ou mapeado a partir de $L_i [b, t]$, conforme ilustrado na Figura 14b, porém de maneira independente para cada banda b . A intensidade acústica multibanda de entrada $L_i [b, t]$ é uma função da intensidade acústica específica do sinal de áudio de entrada $N[b, t]$. Embora possa ser igual à intensidade acústica específica do sinal de áudio de entrada, ela pode ser uma versão suavizada temporalmente e/ou suavizada em frequência da intensidade acústica específica do sinal de áudio.

[00137] A maneira mais direta de se calcular $L_i [b, t]$ é definir a mesma igual à intensidade acústica específica $N[b, t]$. Neste caso, o DRV é feito independentemente em cada banda no banco de filtros auditivos do modelo de intensidade acústica perceptiva ao invés de ser feito de acordo com a mesma entrada versus a razão de intensidade acústica de saída para todas as bandas, como descrito acima intitulada "Função Invariante na frequência e variante no tempo adequada

para ganho automático e controle de faixa dinâmica". Em uma modalidade prática que emprega 40 bandas, o espaçamento destas bandas ao longo do eixo de frequência é relativamente fino a fim de prover uma medição precisa da intensidade acústica. No entanto, a aplicação de um fator de escala DRC independentemente à cada banda pode fazer com que o áudio processado pareça "partido". Para evitar este problema, pode-se calcular $L_i[b, t]$ por meio da suavização da intensidade acústica específica $N[b, t]$ através das bandas de modo que a quantidade de DRC aplicada de uma banda para a outra não varie drasticamente. Isto pode ser obtido ao se definir um filtro de suavização de banda $Q(b)$ e em seguida suavizar a intensidade acústica específica em todas as bandas c de acordo com a soma de convolução padrão:

$$L_i[b, t] = \sum_c Q(b - c) N[c, t], \quad (19)$$

em que $N[c, t]$ é a intensidade acústica específica do sinal de áudio e $Q(b - c)$ é a resposta deslocada na banda do filtro de suavização. A Figura 15 ilustra um exemplo de tal filtro de suavização de banda.

[00138] Quando a função DRC que calcula $L_i[b, t]$ como uma função de $L_o[b, t]$ é fixada para cada banda b , o tipo de mudança feita em cada banda da intensidade acústica específica $N[b, t]$ irá variar dependendo do espectro do áudio que é processado, mesmo que a intensidade acústica geral do sinal continue a mesma. Por exemplo, um sinal de áudio com baixo alto e agudo baixo pode ter o baixo cortado e o agudo reforçado. Um sinal com baixo silencioso e agudo alto pode ocasionar o oposto. O efeito líquido é uma mudança no timbre ou no espectro percebido do áudio, e isto pode ser desejável em certas aplicações.

[00139] No entanto, pode-se desejar fazer um DRC multibanda sem modificar o espectro percebido médio do áudio. Pode-se querer que a modificação média de cada banda seja grosso modo igual e ainda as-

sim permitir que as variações de curto prazo das modificações operem independentemente entre bandas. O efeito desejado pode ser obtido ao fazer com que o comportamento médio do DRC em cada banda fique igual ao de um comportamento de referência. Pode-se escolher este comportamento de referência como o controle DRC desejado para a intensidade acústica de entrada de banda larga $L_i [b, t]$. Sendo assim, a função $L_o [t] = \text{DRC}\{L_i [t]\}$ representa uma versão mediada no tempo da intensidade acústica de entrada de banda larga, e $L_i [b, t]$ representa uma versão mediada no tempo da intensidade acústica de entrada multibanda $L_i [b, t]$. A intensidade acústica de saída multibanda pode então ser calculada como:

$$L_o[b, t] = \frac{\bar{L}_i[b, t]}{\bar{L}_i[t]} \text{DRC}\left\{\frac{\bar{L}_i[t]}{\bar{L}_i[b, t]} L_i[b, t]\right\}. \quad (20)$$

[00140] Observa-se que a intensidade acústica de entrada multibanda é primeiramente graduada de modo a ficar na mesma faixa média que a intensidade acústica de entrada de banda larga. A função DRC concebida para a intensidade acústica de banda larga é então aplicada. Por último, o resultado é graduado de volta à faixa média da intensidade acústica multibanda. Com esta formulação de DRC multibanda, os benefícios de um bombeamento espectral reduzido são mantidos, ao mesmo tempo preservando o espectro percebido médio do áudio.

Função Variante na frequência e Variante no tempo Adequada para Equalização Dinâmica

[00141] Uma outra aplicação de aspectos da presente invenção é a transformação não intencional do espectro percebido variante no tempo do áudio em um espectro percebido invariante no tempo alvo, ao mesmo tempo preservando a faixa dinâmica original do áudio. Pode-se referir a este processamento como uma Equalização Dinâmica (DEQ). Com a equalização tradicional, uma simples filtragem fixa é aplicada ao áudio a fim de mudar o seu espectro. Por exemplo, pode-se aplicar

um baixo fixo ou um reforço de agudo. Este processamento não leva em conta o espectro corrente do áudio e pode, portanto, ser impróprio para alguns sinais, ou seja, para os sinais que já contêm uma quantidade relativamente grande de baixo e agudo. Com o DEQ, o espectro do sinal é medido e o sinal é em seguida dinamicamente modificado a fim de transformar o espectro medido em uma forma desejada essencialmente estática. Para os aspectos da presente invenção, tal forma desejada é especificada nas bandas do banco de filtros e referida como EQ [b]. Em uma modalidade prática, o espectro medido deve representar a forma espectral média do áudio que pode ser gerada por meio da suavização da intensidade acústica específica $N[b, t]$ através do tempo. Pode-se referir à intensidade acústica específica suavizada como $\bar{N}[b, t]$. De acordo com o controle DRC multibanda, pode-se não querer que a modificação DEQ varie drasticamente de uma banda para a outra, e, sendo assim, uma função de suavização de banda pode ser aplicada de modo a gerar um espectro suavizado na banda $L[b, t]$:

$$\bar{L}[b, t] = \sum_c Q(b - c) \bar{N}[c, t]. \quad (21)$$

[00142] A fim de preservar a faixa dinâmica original do áudio, o espectro desejado EQ [b] deve ser normalizado de modo a apresentar a mesma intensidade acústica geral que a forma espectral medida dada por $L[b, t]$. Pode-se referir a esta forma espectral normalizada como $L_{EQ}[b, t]$:

$$\bar{L}_{EQ}[b, t] = \left(\frac{\sum_c \bar{L}[c, t]}{\sum_c EQ[c]} \right) EQ[b]. \quad (22)$$

[00143] Finalmente, o parâmetro específico é vulgarmente calculado como:

$$\hat{N}[b, t] = \left(\frac{\bar{L}[b, t]}{\bar{L}_{EQ}[b, t]} \right)^{1-\beta} \frac{\bar{L}_{EQ}[b, t]}{\bar{L}[b, t]} N[b, t], \quad (23)$$

em que β é um parâmetro especificado pelo usuário variando de zero a um, indicando o grau de DEQ a ser aplicado. Olhando a Equação 23,

nota-se que, quando $\beta = 0$, a intensidade acústica específica original não é modificada, e quando $\beta = 1$, a intensidade acústica específica é graduada pela razão espectral desejada para a forma espectral medida.

[00144] Uma maneira conveniente de gerar a forma espectral desejada EQ[b] é o usuário definir a mesma igual a $\bar{L}[b, t]$ conforme medida para a mesma peça de áudio cujo equilíbrio espectral o usuário considera agradável. Em uma modalidade prática, por exemplo, conforme mostrado na Figura 16, o usuário pode ser provido com um botão ou outro atuador adequado 507 que, quando atuado, provoca a captura da medida corrente da forma espectral do áudio $L[b, t]$ e, em seguida, armazena esta medida como uma predefinição (em Captura e Armazenamento Predefinidos de Intensidade acústica alvo específica 506) que poderá ser mais tarde carregada para o EQ [b] quando o DEQ estiver habilitado (como pela seleção predefinida 508). A Figura 16 é uma versão simplificada da Figura 7 na qual apenas uma única linha é mostrada para representar múltiplas bandas do Banco de filtros de análise 100 para o Banco de filtros de síntese 110. O exemplo da Figura 17 também provê uma Modificação de intensidade acústica específica de EQ Dinâmico 505 que provê uma modificação à intensidade acústica específica medida por meio da função ou dispositivo 104 de acordo com a equalização dinâmica, conforme explicado acima.

Processamento Combinado

[00145] Pode-se querer combinar todos o processamento acima descrito, incluindo o Controle de volume (VC), o controle AGC, o DRC, e o DEQ, em um único sistema. Uma vez que cada um destes processos pode ser representado como uma graduação da intensidade acústica específica, todos eles são facilmente combinados, como segue:

$$\hat{N}[b, t] = (\Xi_{VC}[b, t] \Xi_{AGC}[b, t] \Xi_{DRC}[b, t] \Xi_{DEQ}[b, t]) N[b, t], \quad (24)$$

em que $\Xi[b, t]$ representa os fatores de escala associados ao processo

""". Um único conjunto de ganhos $G[b, t]$ pode então ser calculado para a intensidade acústica alvo específica que representa o processamento combinado.

[00146] Em alguns casos, os fatores de escala de um processo ou combinação de processos de modificação de intensidade acústica podem flutuar rápido demais ao longo do tempo e produzir artefatos no áudio processado resultante. Pode, portanto, ser desejável suavizar um subconjunto destes fatores de escala. Em geral, os fatores de escala de VC e DEQ variam suavemente ao longo do tempo, mas a suavização da combinação dos fatores de escala de AGC e DRC pode ser necessária. A combinação destes fatores de escala é representada por:

$$\Xi_c[b, t] = \Xi_{AGC}[b, t] \Xi_{DRC}[b, t] \quad (25)$$

[00147] A noção básica atrás da suavização é que os fatores de escala combinados devem reagir rapidamente quando a intensidade acústica específica aumenta, e que os fatores de escala devem ser mais fortemente suavizados quando a intensidade acústica específica diminui. Esta noção corresponde à prática bem conhecida de se utilizar um rápido ataque e uma lenta liberação na concepção dos compressores de áudio. As constantes de tempo apropriadas para a suavização dos fatores de escala podem ser calculadas ao se suavizar ao longo do tempo uma versão suavizada na banda da intensidade acústica específica. Primeiramente, uma versão suavizada na banda da intensidade acústica específica é computada:

$$L[b, t] = \sum_c Q(b - c) N[c, t], \quad (26)$$

em que $N[b, t]$ é a intensidade acústica específica do sinal de áudio e $Q(b - c)$ é a resposta deslocada na banda do filtro de suavização, como na Equação 19 acima.

[00148] A versão suavizada no tempo desta intensidade acústica específica suavizada na banda é então calculada como:

$$\bar{L}[b, t] = \lambda[b, t]L[b, t] + (1 - \lambda[b, t])\bar{L}[b, t-1] \quad (27)$$

em que o coeficiente de suavização dependente de banda $\lambda [b, t]$ é dado por:

$$\lambda[b, t] = \begin{cases} \lambda_{fast}, & L[b, t] > \bar{L}[b, t] \\ \lambda_{slow}, & L[b, t] \leq \bar{L}[b, t] \end{cases} \quad (28)$$

[00149] Os fatores de escala combinados suavizados são em seguida calculados como:

$$\bar{\Xi}_c[b, t] = \lambda_M[b, t]\Xi_c[b, t] + (1 - \lambda_M[b, t])\bar{\Xi}_c[b, t-1], \quad (29)$$

em que $\lambda_M [b, t]$ é uma versão suavizada na banda de $\lambda [b, t]$:

$$\lambda_M[b, t] = \left(\frac{1}{\sum_c Q(c)} \right) \sum_c Q(b-c) \lambda[b, t]. \quad (30)$$

[00150] A suavização de banda do coeficiente de suavização impede que os fatores de escala suavizados no tempo mudem drasticamente através das bandas. A suavização na banda e no tempo dos fatores de escala descritos resulta em um áudio processado contendo menos artefatos perceptivos objetáveis.

Compensação de Ruído

[00151] Em muitos ambientes de reprodução de áudio existe ruído de fundo que interfere no áudio que um ouvinte deseja escutar. Por exemplo, um ouvinte em um automóvel em movimento pode estar ouvindo música de um sistema estéreo instalado e o barulho do motor e da rua pode significativamente alterar a percepção da música. Em particular, para as partes do espectro nas quais a energia do ruído é significativa com relação à energia da música, a intensidade acústica percebida da música é reduzida. Quando o nível do ruído é muito grande, a música fica completamente mascarada. Com relação a um aspecto da presente invenção, pode-se querer escolher ganhos $G [b, t]$ de modo que a intensidade acústica específica do áudio processado na presença do ruído de interferência seja igual à intensidade acústica alvo específica $N[b, t]$. Para se chegar a este efeito, pode-se utilizar o con-

ceito de intensidade acústica parcial, conforme definido por Moore e Glasberg, supra. Pressupõe-se que se possa obter uma medição do ruído por si só e uma medição do áudio por si só. $E_N[b, t]$ representaria a excitação do ruído e $E_A[b, t]$ representaria a excitação do áudio. A intensidade acústica específica combinada do áudio e do ruído é então dada por:

$$N_{TOT}[b, t] = \Psi\{E_A[b, t] + E_N[b, t]\} \quad (31)$$

em que, mais uma vez, $\Psi\{.\}$ representa a transformação não linear a partir da excitação para a intensidade acústica específica. Pode-se pressupor que a audição de um ouvinte divide a intensidade acústica específica combinada entre a intensidade acústica específica parcial do áudio e a intensidade acústica específica parcial do ruído de maneira a preservar a intensidade acústica específica combinada:

$$N_{TOT}[b, t] = N_A[b, t] + N_N[b, t] \quad (32)$$

[00152] A intensidade acústica específica parcial do áudio, $N_A[b, t]$, é o valor que se quer controlar, e, portanto, deve-se solucionar este valor. A intensidade acústica específica parcial do ruído pode ser aproximada, como:

$$N_N[b, t] = \left(\frac{E_{TN}[b, t]}{E_A[b, t]} \right)^k \left(\Psi\{E_N[b, t] + E_{TN}[b, t]\} - \Psi\{E_{TQ}[b]\} \right) \quad (33)$$

em que $E_{TN}[b, t]$ é o limite mascarado na presença do ruído, $E_{TQ}[b]$ é o limite da audição em silêncio na banda b , e k é um expoente entre zero e um.

[00153] Ao se combinar as Equações 31 a 33, chega-se a uma expressão para a intensidade acústica específica parcial do áudio:

$$N_A[b, t] = \Psi\{E_A[b, t] + E_N[b, t]\} - \left(\frac{E_{TN}[b, t]}{E_A[b, t]} \right)^k \left(\Psi\{E_N[b, t] + E_{TN}[b, t]\} - \Psi\{E_{TQ}[b]\} \right) \quad (34)$$

[00154] Nota-se que quando a excitação do áudio é igual ao limite mascarado do ruído ($E_A[b, t] = E_{TN}[b, t]$), a intensidade acústica específica parcial do áudio será igual à intensidade acústica de um sinal no limite em silêncio, o que vem a ser um resultado desejado. Quando a

excitação do áudio é muito maior que a do ruído, o segundo termo da Equação 34 desaparece, e a intensidade acústica específica do áudio será aproximadamente igual ao que seria se o ruído não estivesse presente. Em outras palavras, uma vez que o áudio fica muito mais alto que o ruído, o ruído é mascarado pelo áudio. O expoente K é escolhido empiricamente de modo a oferecer um bom ajuste aos dados da intensidade acústica de um tom no ruído como uma função da razão sinal para ruído. Moore e Glasberg descobriram que um valor de $k = 0.3$ é apropriado. O limite mascarado do ruído pode ser aproximado como uma função da própria excitação do ruído.

$$E_{TN}[b, t] = K[b]E_N[b, t] + E_{TP}[b] \quad (35)$$

em que $K[b]$ é uma constante que aumenta em bandas de frequência mais baixas. Sendo assim, a intensidade acústica específica parcial do áudio dada pela Equação 34 pode ser representada abstratamente como uma função da excitação do áudio e da excitação do ruído:

$$N_A[b, t] = \Phi\{E_A[b, t], E_N[b, t]\}. \quad (36)$$

[00155] Um solucionador de ganho modificado pode em seguida ser utilizado no sentido de calcular os ganhos $G[b, t]$ de tal modo que a intensidade acústica específica parcial do áudio processado na presença do ruído fique igual à intensidade acústica alvo específica:

$$\hat{N}[b, t] = \Phi\{G^2[b, t]E_A[b, t], E_N[b, t]\} \quad (37)$$

[00156] A Figura 17 ilustra o sistema da Figura 7 com o Solucionador de ganho original 106 substituído pelo Solucionador de ganho de Compensação de Ruído descrito 206 (observe que as múltiplas linhas verticais entre os blocos representando as múltiplas bandas do banco de filtros foram substituídas por uma única linha a fim de simplificar o diagrama). Além disso, a figura ilustra a medição da excitação de ruído (pelo Banco de filtros de análise 200, pelo Filtro de transmissão 201, pela Excitação 202 e pela Suavização 203 de uma maneira correspondente à operação dos blocos 100, 101, 102 e 103) que alimenta o

novo solucionador de ganho 206 juntamente com a excitação do áudio (a partir da Suavização 103) e da intensidade acústica alvo específica (a partir da Modificação de nível SL 105).

[00157] Em seu modo mais básico de operação, a Modificação SL 105 da Figura 17 pode simplesmente definir uma intensidade acústica alvo específica $N[b, t]$ igual à intensidade acústica específica original do áudio $N[b, t]$. ou seja, a Modificação SL provê uma graduação α de fator de escala invariante na frequência da intensidade acústica específica do sinal de áudio, na qual $\alpha = 1$. Com uma disposição como a da Figura 17, os ganhos são calculados de modo que o espectro da intensidade acústica percebida do áudio processado na presença do ruído seja igual aos do espectro de intensidade acústica do áudio na ausência do ruído. Além disso, qualquer técnica ou combinação de técnicas previamente descritas para a computação da intensidade acústica alvo específica como uma função do original, incluindo VC, AGC, DRC, e DEQ, pode ser utilizada em conjunto com o sistema de modificação de intensidade acústica de compensação de ruído.

[00158] Em uma modalidade prática, a medição do ruído pode ser feita a partir de um microfone colocado no ou próximo do ambiente no qual o áudio será tocado. De maneira alternativa, um conjunto predeterminado de excitações de ruído padrão pode ser utilizado próximo do espectro de ruído antecipado de acordo com diversas condições. Por exemplo, o ruído de uma cabine de automóvel pode ser pré-analisado a várias velocidades de direção e em seguida armazenado como uma tabela de pesquisa da excitação de ruído versus velocidade. A excitação de ruído alimentada para o Solucionador de ganho 206 da Figura 17 pode ser aproximada desta tabela de pesquisa à medida que a velocidade do automóvel varia.

Aproximações à Intensidade acústica específica

[00159] Embora a presente invenção funcione melhor quando uma

medida precisa da intensidade acústica específica é usada, algumas aplicações podem requerer o uso de uma aproximação mais grosseira a fim de reduzir a complexidade computacional. Com uma aproximação adequada, uma estimativa aceitável e modificação da intensidade acústica percebida poderão ainda ser obtidas. Tal aproximação deve tentar preservar, pelo menos em parte, vários aspectos chave da percepção de intensidade acústica. Primeiramente, a aproximação deve pelo menos capturar grosso modo a sensibilidade variante na percepção de intensidade acústica com frequência. Em termos gerais, a aproximação deve exibir menor sensibilidade a frequências mais baixas e mais altas em comparação às frequências intermediárias. Em segundo lugar, a aproximação deve exibir um crescimento não linear da intensidade acústica com nível de sinal. Especificamente, o crescimento da intensidade acústica específica deve ser mais rápido para os sinais de nível baixo próximos do limite de audição e então diminuir a uma taxa de crescimento constante à medida que o nível de sinal aumenta. por último, a aproximação deve exibir a propriedade de somar intensidade acústica, isto querendo dizer que, para um nível de sinal constante, a intensidade acústica total (a integração da intensidade acústica específica através da frequência) aumentará à medida que a largura de banda de um sinal aumenta.

[00160] Uma maneira de se reduzir a complexidade da computação da intensidade acústica específica e ao mesmo tempo manter as propriedades desejadas de percepção de intensidade acústica é usar um banco de filtros com menos bandas, no qual as bandas não podem ser espaçadas uniformemente em uma escala de taxa de banda crítica. Por exemplo, pode-se usar um banco de filtros de 5 bandas com as bandas distribuídas uniformemente em uma escala de frequência linear em oposição às 40 bandas descritas anteriormente. Existem muitas técnicas para uma eficiente implementação de tais bancos de filtros,

por exemplo, os bancos de filtros modulados em co-seno (P. P. Vaidyanthan, *MultiRate Systems and Filter Banks*, 1993, Prentice Hall). Como um exemplo geral, considere-se um banco de filtros com bandas B , no qual cada banda é descrita por uma resposta de impulso de domínio de tempo $h_b[n]$. Suponha, ainda, que o banco de filtros seja uma reconstrução quase perfeita, isto significando que:

$$x[n] \cong \sum_{b=1}^B h_b[n] * x[n] \quad (38)$$

[00161] A resposta de frequência de cada banda b pode ser caracterizada por uma frequência central f_b e uma largura de banda Δf_b em Hertz. A frequência central equivalente e a largura de banda em unidades de ERB são então obtidas por:

$$e_b = HzToERB(f_b) \quad (39a)$$

$$\Delta e_b = HzToERB\left(f_b + \frac{\Delta f_b}{2}\right) - HzToERB\left(f_b - \frac{\Delta f_b}{2}\right) \quad (39b)$$

[00162] Quando o número de bandas é relativamente pequeno, a largura de banda e_b de cada banda será provavelmente maior que 1 ERB.

[00163] Quando $x_b[n] = h_b[n] * [n]$ representa o sinal de áudio associado à cada banda, um sinal de excitação suavizado $E[b, n]$ poderá ser computado de maneira similar às Equações 7 e 8 por meio da computação da força rms suavizada de $x_b[n]$ ponderada pela resposta de frequência do filtro de transmissão amostrado na frequência f_b e o inverso da largura de banda em unidades de ERB:

$$\bar{E}[b, n] = \lambda_b \bar{E}[b, n] + (1 - \lambda_b) \frac{|P(f_b)|^2}{e_b} x_b^2[n] \quad (40)$$

[00164] A ponderação da excitação da banda b por $1/e_b$ efetivamente distribui a energia dentro daquela banda uniformemente por todas as bandas críticas sub-somadas dentro da mesma. Pode-se, alternativamente, atribuir toda a energia à banda crítica, cuja frequência central fica mais próxima da frequência central f_b da banda, porém a distribui-

ção uniforme da energia oferece uma melhor aproximação para a maioria dos sinais de áudio do mundo real.

[00165] Com a excitação $E[b, n]$, pode-se continuar a computar a intensidade acústica específica $N[b, n]$, a intensidade acústica alvo específica $N[b, n]$, e os ganhos $G[b, n]$, conforme apresentado acima, com exceção de que, aqui, o número total de bandas poderá ser muito menor, desta forma reduzindo a complexidade. Ainda, uma modificação precisa ser feita na computação da intensidade acústica geral $L[n]$: a intensidade acústica específica precisa ser ponderada pelo número de bandas ERB dentro da banda b quando a mesma é somada através das bandas:

$$L[n] = \sum_b e_b N[b, n] \quad (41)$$

[00166] Finalmente, um sinal modificado $y[n]$ pode ser produzido ao se somar cada um dos sinais de banda ponderados pelo ganho da respectiva banda:

$$y[n] = \sum_b G[b, n] x_b[n] \quad (42)$$

[00167] Para simplificar a exposição, as computações acima são mostradas para serem realizadas para cada período de amostra n do sinal $x[n]$. Na prática, no entanto, a excitação pode ser subamostrada a uma taxa muito menor através do tempo, e, em seguida, todo o processamento de intensidade acústica subsequente pode ser feito nesta taxa reduzida. Quando os ganhos são aplicados no fim, os mesmos poderão então ser supermostrados através de interpolação antes de serem aplicados aos sinais de banda.

[00168] Acima encontra-se tão-somente um exemplo de uma aproximação mais grosseira a uma intensidade acústica específica apropriada para a presente invenção. Outras aproximações são possíveis, e a presente invenção tem a intenção de incluir o uso de todas estas aproximações.

Implementação

[00169] A presente invenção pode ser implementada em hardware ou software, ou em uma combinação de ambos (por exemplo, arrays lógicos programáveis). A menos que de outra forma especificado, os algoritmos incluídos como parte da presente invenção não se referem inerentemente a nenhum computador em particular ou outro aparelho. A este aspecto, várias máquinas de uso geral poderão ser usadas com os programas descritos de acordo com os ensinamentos do presente documento, ou poderá ser mais conveniente construir aparelhos mais especializados (por exemplo, circuitos integrados) a fim de executar as etapas de método requeridas. Sendo assim, a presente invenção pode ser implementada em um ou mais programas de computador que executem um ou mais sistemas computacionais programáveis, cada qual compreendendo pelo menos um processador, pelo menos um sistema de armazenamento de dados (incluindo as memórias voláteis e/ou não voláteis e/ou elementos de armazenamento), pelo menos um dispositivo ou porta de entrada, e pelo menos um dispositivo ou porta de saída. Um código de programa é aplicado aos dados de entrada a fim de realizar as funções descritas no presente documento e gerar as informações de saída. As informações de saída são aplicadas a um ou mais dispositivos de saída, de uma maneira conhecida.

[00170] Cada programa pode ser implementado em qualquer linguagem computacional desejada (incluindo máquina, montagem, ou linguagens de programação procedimental de alto nível, lógica, ou orientada a objeto) no sentido de se comunicar com um sistema de computador. De qualquer maneira, a linguagem deve ser uma linguagem compilada ou interpretada.

[00171] Cada programa de computador deve, de preferência, ser armazenado em ou transferido para um meio ou dispositivo de armazenamento (por exemplo, uma memória ou meio de estado sólido, ou

um meio magnético ou óptico) legível por um computador programável de uso geral ou de uso especial, para a configuração e operação do computador quando o meio ou dispositivo de armazenamento é lido por um sistema de computador a fim de realizar os procedimentos descritos no presente documento. O sistema inventivo pode também ser considerado de modo a ser implementado como um meio de armazenamento legível em computador, configurado com um programa de computador, no qual o meio de armazenamento assim configurado faz com que o sistema computacional opere de uma maneira específica e predefinida a fim de executar as funções descritas no presente documento.

[00172] Foram descritas diversas modalidades da presente invenção. No entanto, deve-se entender que várias modificações podem ser feitas sem se afastar do espírito e âmbito de aplicação da presente invenção. Por exemplo, algumas etapas aqui descritas podem ser independentes de ordem, e, assim, poderão ser realizadas em uma ordem diferente da descrita.

REIVINDICAÇÕES

1. Método para controlar uma característica de intensidade acústica particular de um sinal de áudio, **caracterizado pelo fato de que** a característica de intensidade acústica particular é, ou uma aproximação da intensidade acústica específica, que é uma medida de intensidade acústica perceptiva como uma função da frequência e do tempo, ou uma aproximação da intensidade acústica específica parcial, que é uma medida de intensidade acústica perceptiva do sinal na presença de um sinal de interferência secundário como uma função da frequência e do tempo, o método compreendendo as etapas de:

calcular, como uma função do sinal de áudio, uma aproximação a uma intensidade acústica alvo específica, em que a aproximação a uma intensidade acústica alvo específica é obtida com o uso de um banco de filtros no qual há menos bandas do que seriam obtidas com o uso de um banco de filtros no qual as bandas foram uniformemente espaçadas na escala ERB com um espaçamento ERB desejado de 1,

derivar parâmetros de modificação utilizáveis para modificar o sinal de áudio de uma forma variante no tempo e/ou na frequência a fim de reduzir a diferença entre sua a característica de intensidade acústica particular e a aproximação para uma intensidade acústica alvo específica, e

a) aplicar os parâmetros de modificação ao sinal de áudio, a fim de reduzir a diferença entre a sua característica de intensidade acústica particular e a aproximação para uma intensidade acústica alvo específica, ou

b) transmitir ou armazenar os parâmetros de modificação e o sinal de áudio para aplicação temporalmente separada e/ou espacialmente separada dos parâmetros de modificação para o sinal de áudio reduzir a diferença entre a sua característica de intensidade acústi-

ca particular e a aproximação para uma intensidade acústica alvo específica.

2. Método, de acordo com a reivindicação 1, **caracterizado pelo fato de que** cada banda no banco de filtros é maior do que 1 ERB.

3. Método, de acordo com a reivindicação 1 ou 2, **caracterizado pelo fato de que** as bandas no banco de filtros são espaçadas de modo não-uniforme em uma escala de taxa de banda crítica.

4. Aparelho, **caracterizado pelo fato de que** é adaptado para executar todas as etapas do método conforme definido em qualquer uma das reivindicações 1 a 3.

5. Método para controlar uma característica de intensidade acústica particular de um sinal de áudio, **caracterizado pelo fato de que** a característica de intensidade acústica particular é ou uma aproximação da intensidade acústica específica, que é uma medida de intensidade acústica perceptiva como uma função da frequência e do tempo, ou uma aproximação da intensidade acústica parcial, que é uma medida de intensidade acústica perceptiva do sinal de áudio na presença de um sinal de interferência secundário como uma função da frequência e do tempo, o método compreendendo as etapas de:

receber o sinal de áudio a partir de uma transmissão ou reprodução de um meio de armazenamento, e

a) parâmetros de modificação para modificar o sinal de áudio, os parâmetros de modificação tendo sido calculados a partir de uma aproximação para uma intensidade acústica alvo específica que emprega um banco de filtros no qual há menos bandas do que seriam obtidas com o uso de um banco de filtros no qual as bandas foram uniformemente espaçadas na escala ERB com um espaçamento ERB desejado de 1, a aproximação tendo sido calculada como uma função do sinal de áudio, ou

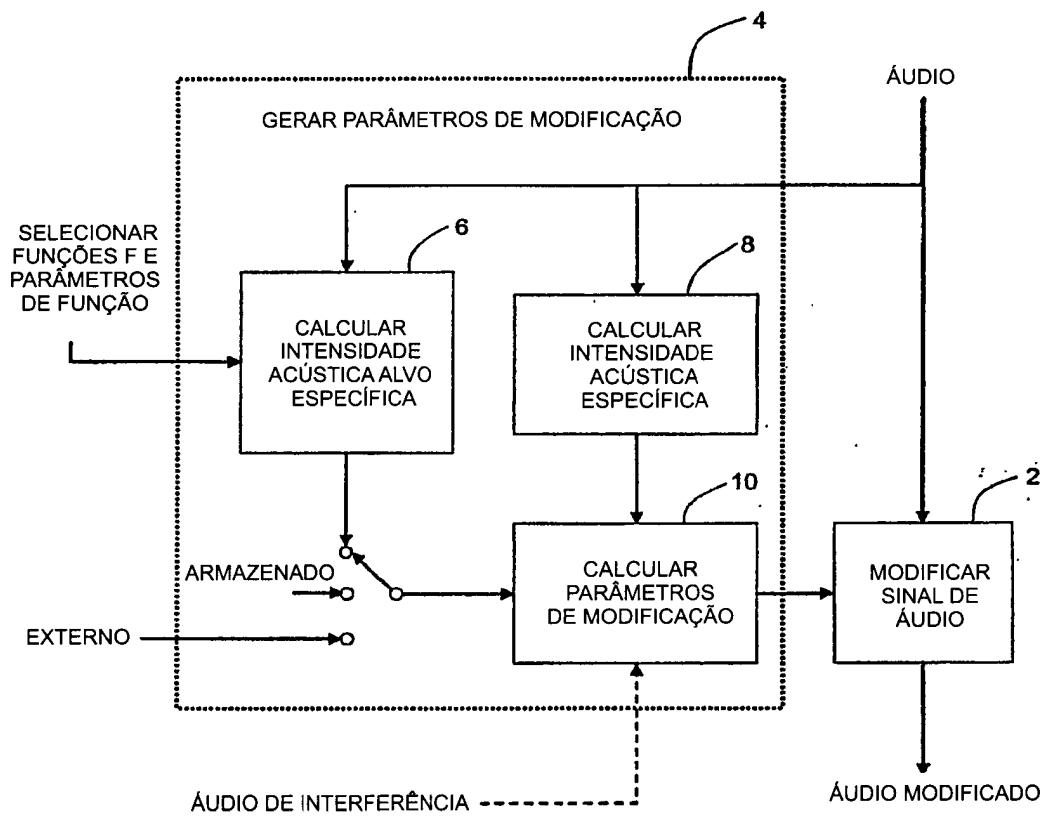
b) uma aproximação para uma intensidade acústica alvo específica que emprega um banco de filtros no qual há menos bandas do que seriam obtidas com o uso de um banco de filtros no qual as bandas foram uniformemente espaçadas na escala ERB com um espaçamento ERB desejado de 1, ou uma representação da aproximação para aquela intensidade acústica alvo específica, a aproximação tendo sido calculada como uma função do sinal de áudio, e

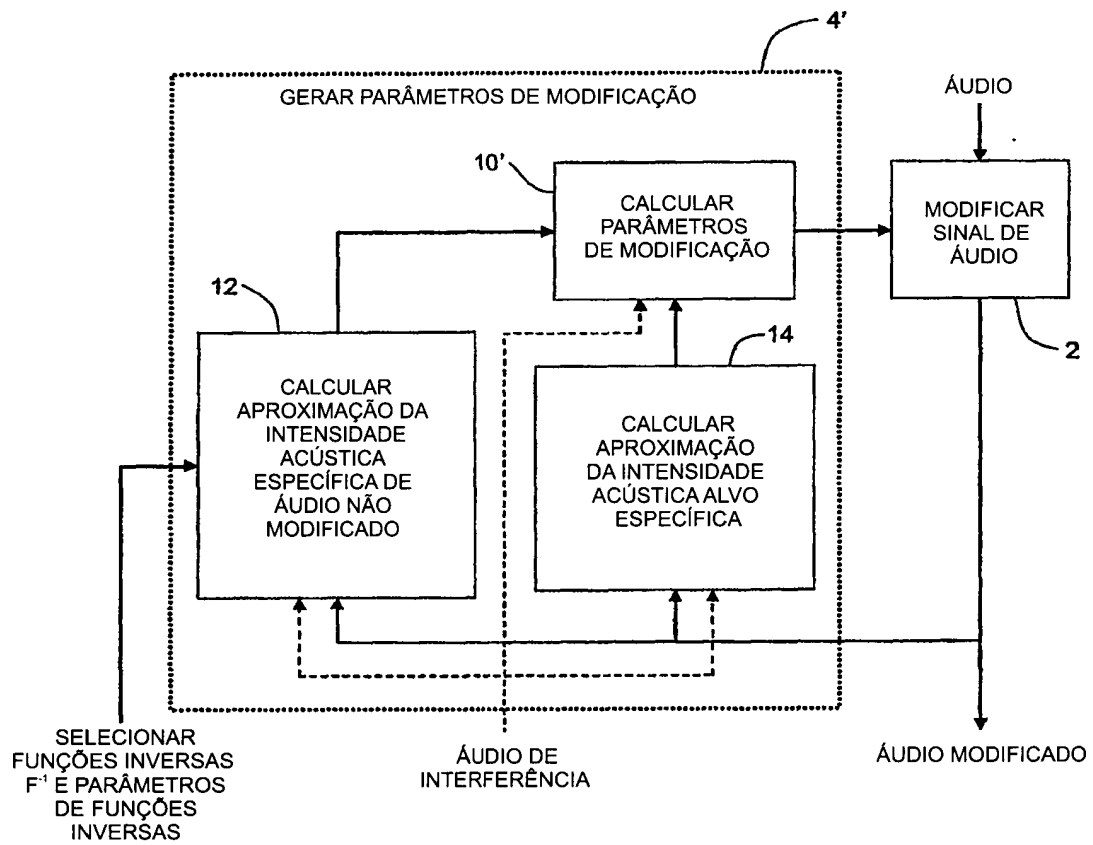
modificar, de uma forma variante no tempo e/ou na frequência, o sinal de áudio em resposta a a) os parâmetros de modificação recebidos, ou b) parâmetros de modificação calculados a partir da aproximação para uma intensidade acústica alvo específica ou sua representação de modo a reduzir a diferença entre a sua característica de intensidade acústica particular do sinal de áudio e a aproximação para uma intensidade acústica alvo específica.

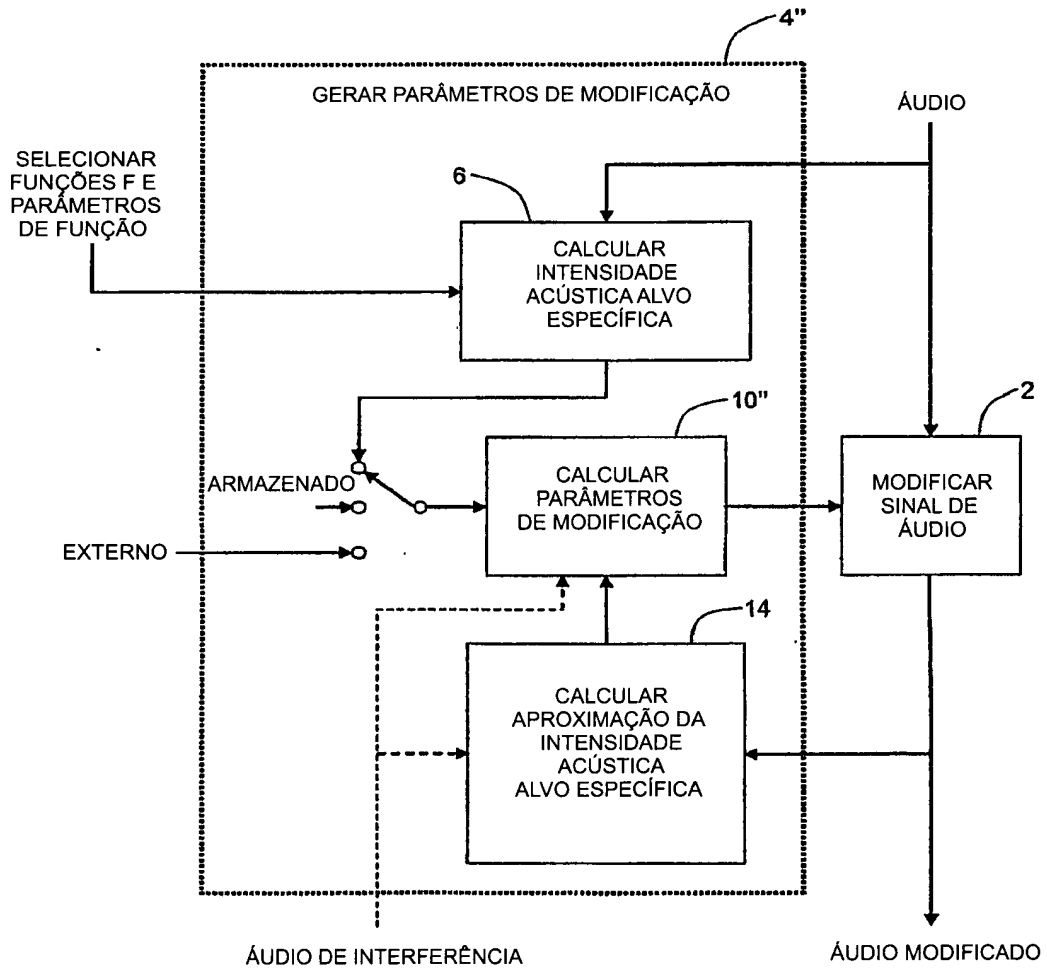
6. Método, de acordo com a reivindicação 5, **caracterizado pelo fato de que** cada banda no banco de filtros que é empregado na aproximação para a intensidade acústica alvo específica é maior do que 1 ERB.

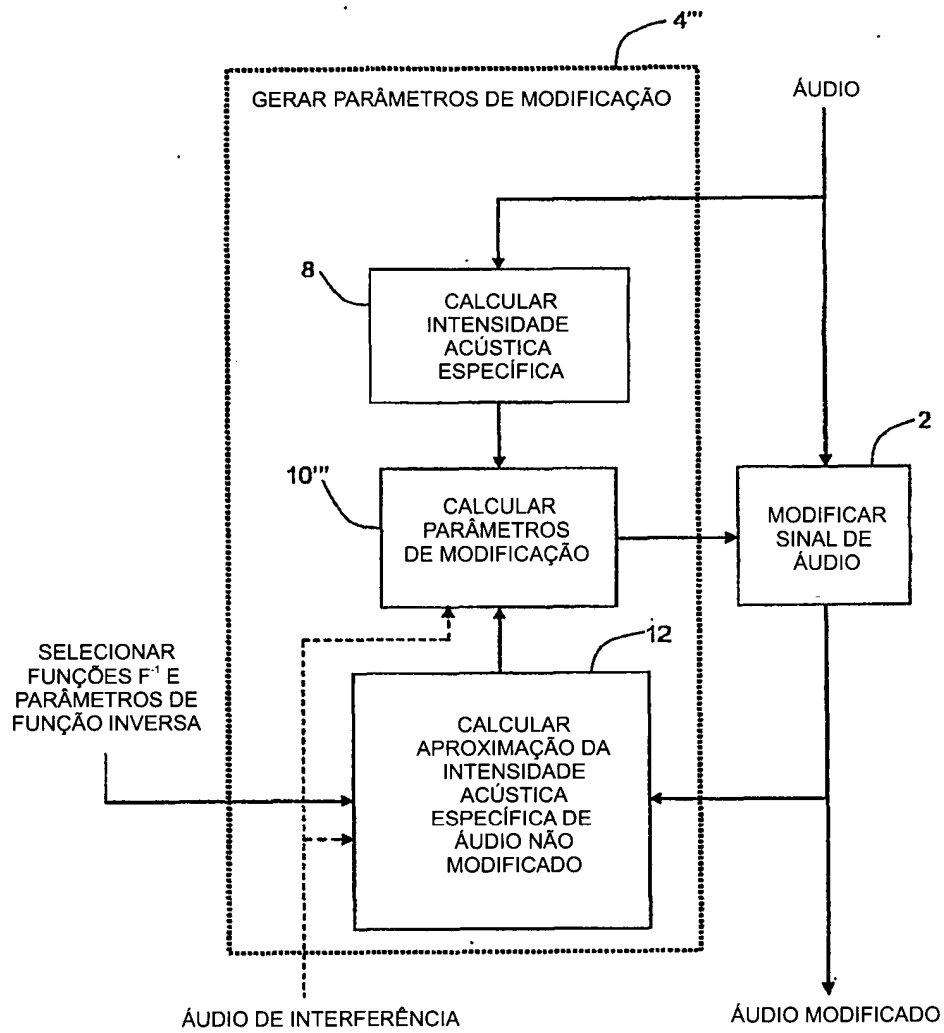
7. Método, de acordo com a reivindicação 5 ou 6, **caracterizado pelo fato de que** as bandas no banco de filtros são espaçadas de modo não-uniforme em uma escala de taxa de banda crítica.

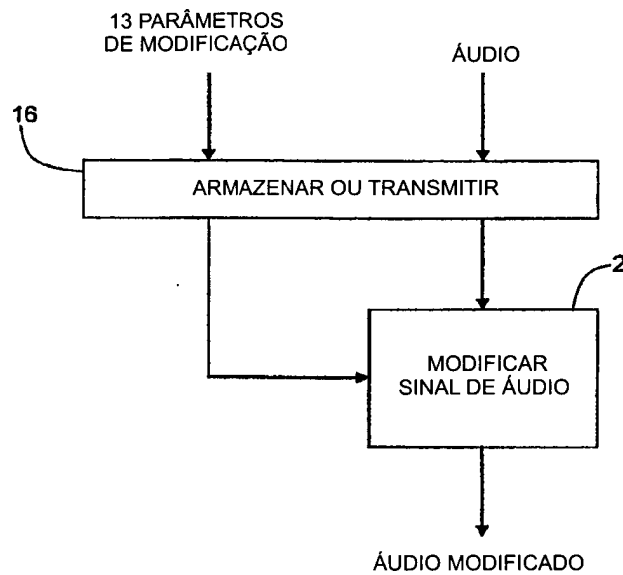
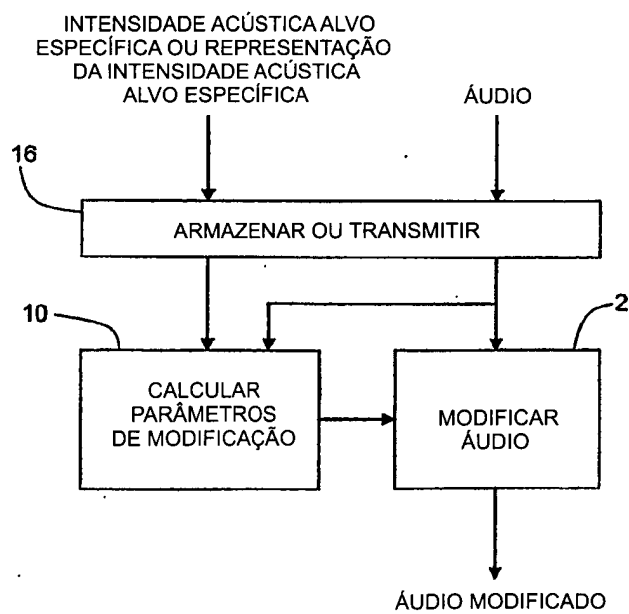
8. Aparelho, **caracterizado pelo fato de que** é adaptado para executar todas as etapas do método conforme definido em qualquer uma das reivindicações 5 a 7.

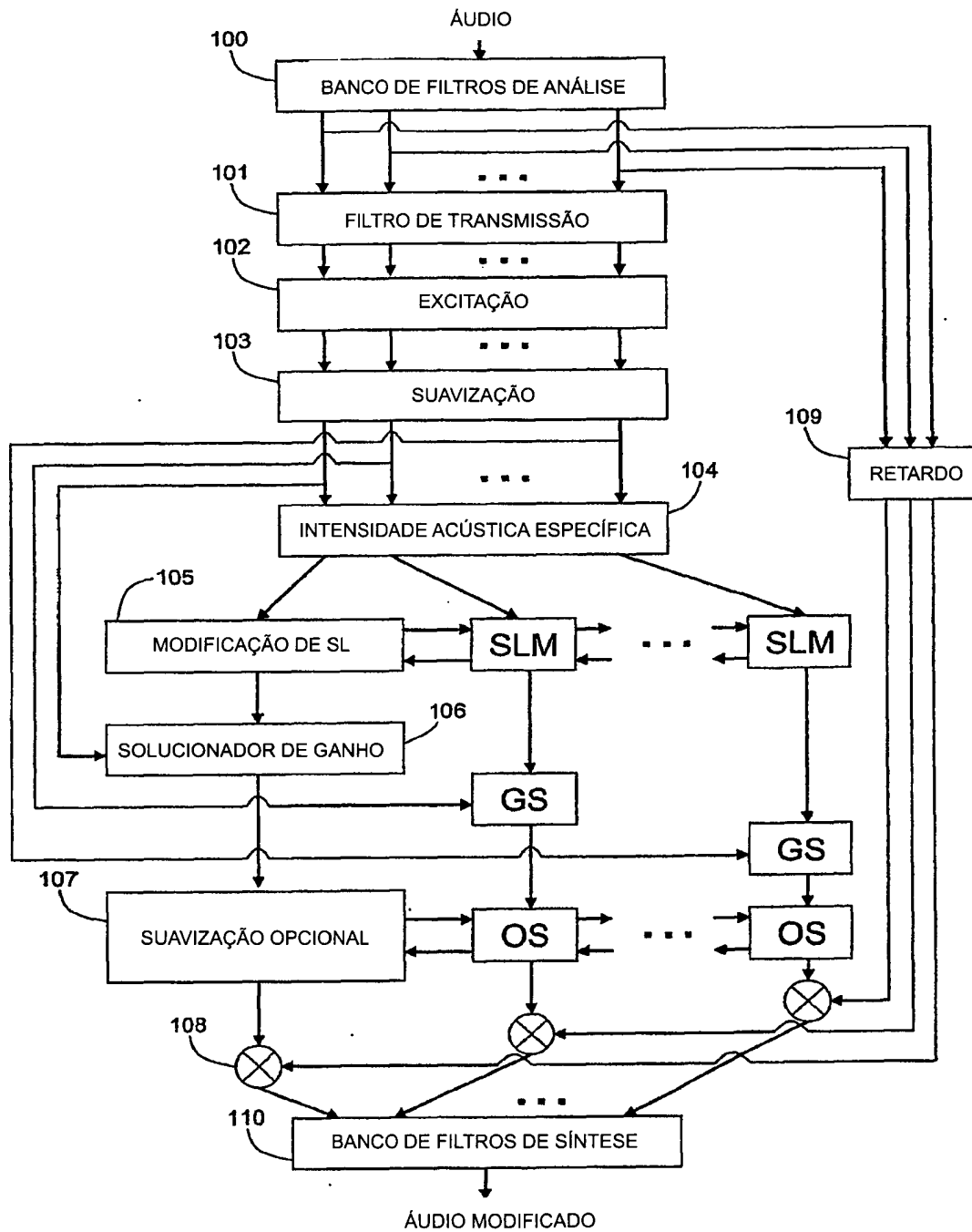
**FIG. 1**

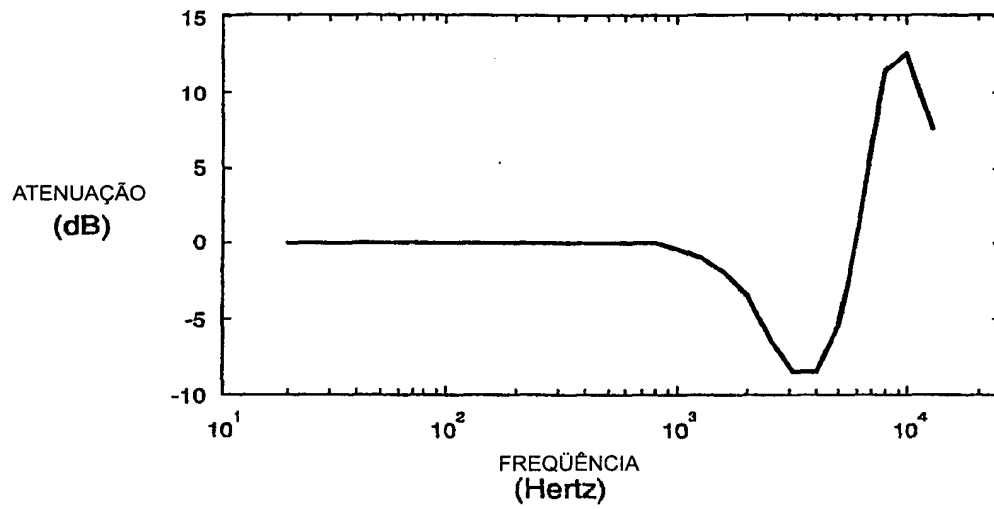
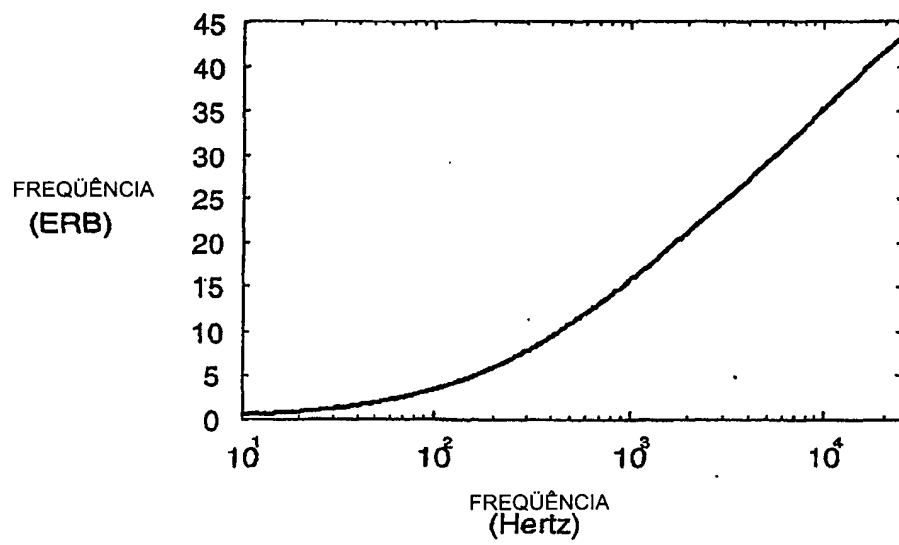
**FIG. 2**

**FIG. 3**

**FIG. 4**

**FIG. 5****FIG. 6**

**FIG. 7**

**FIG. 8****FIG. 9**

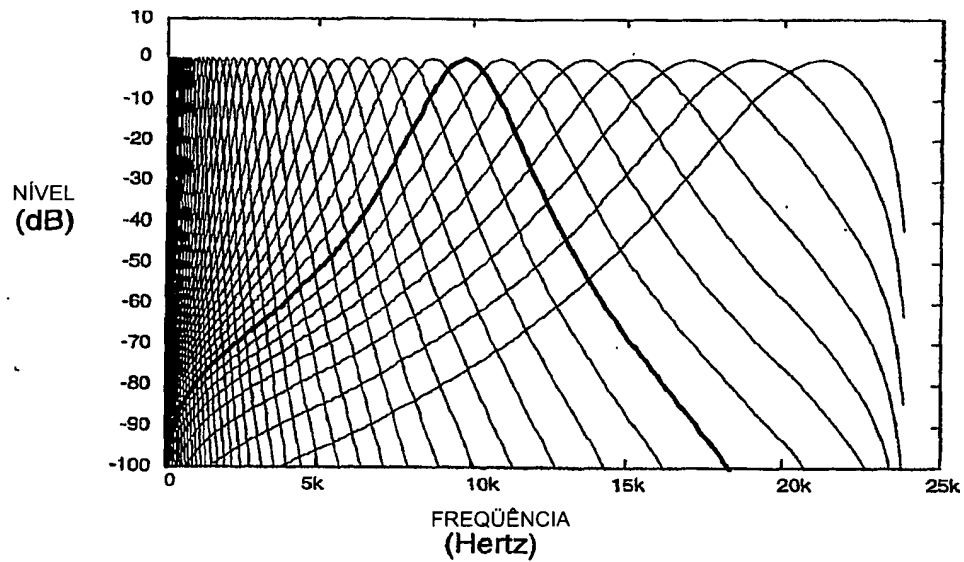


FIG. 10

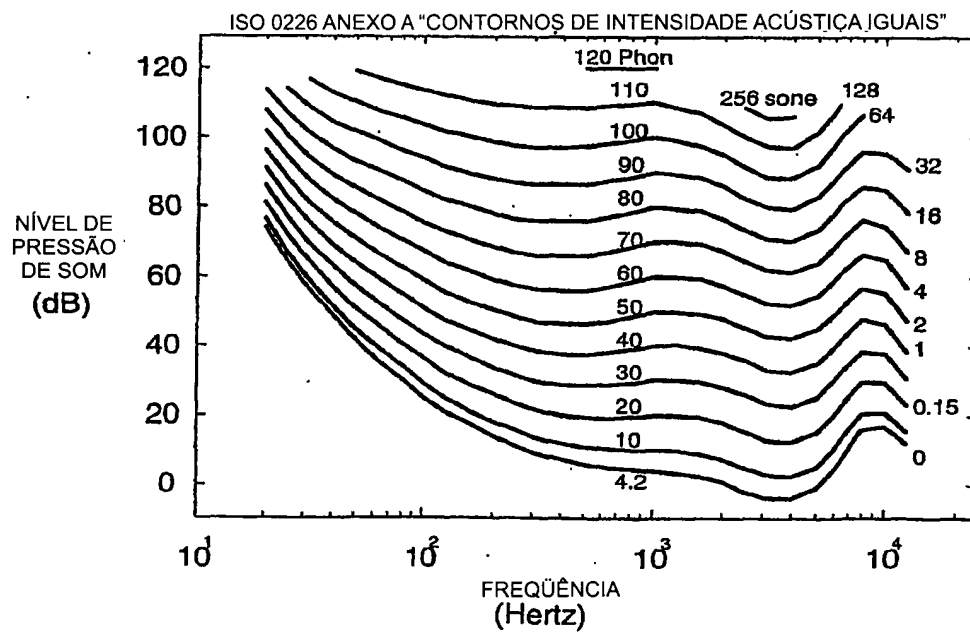


FIG. 11

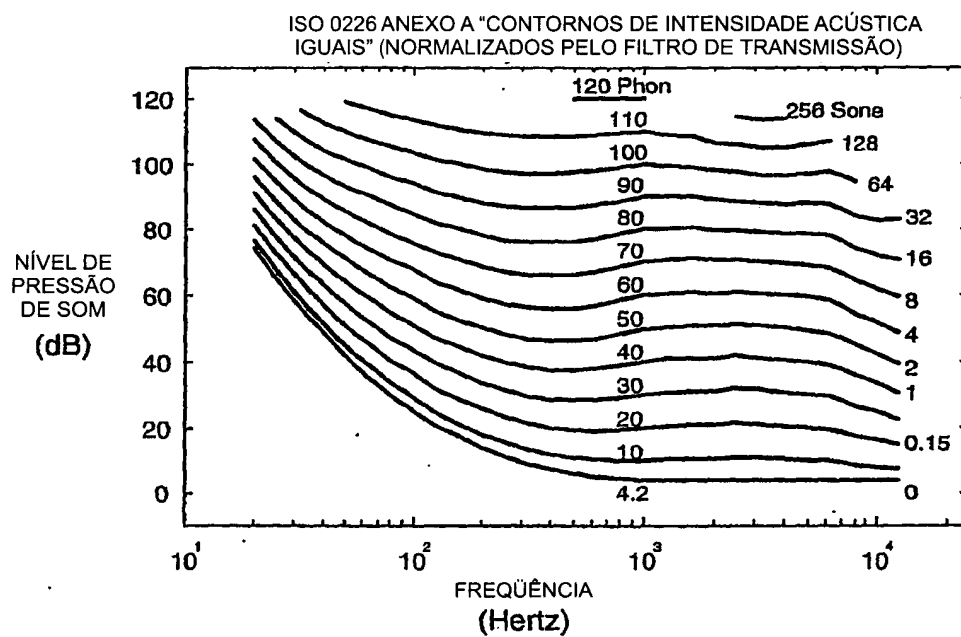


FIG. 12

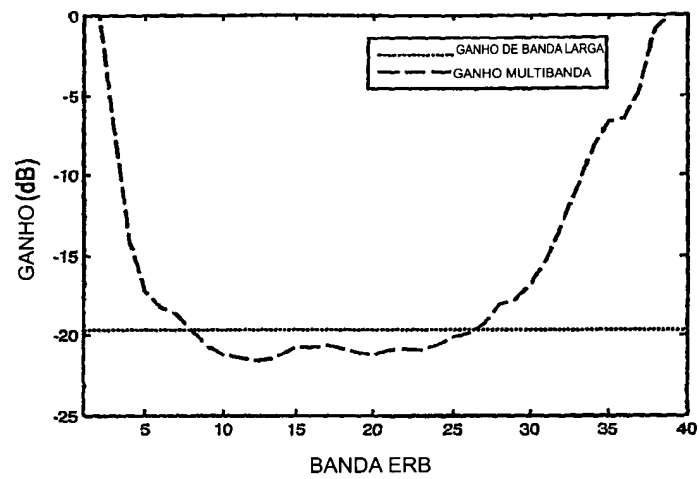
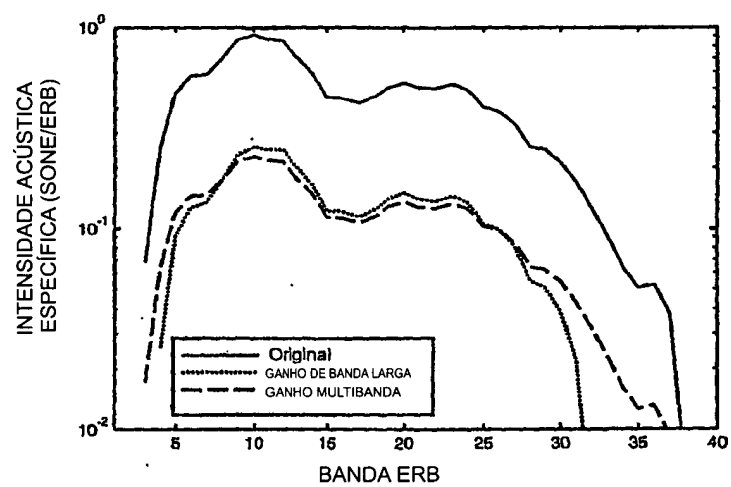
**FIG. 13a****FIG. 13b**

FIG. 14a

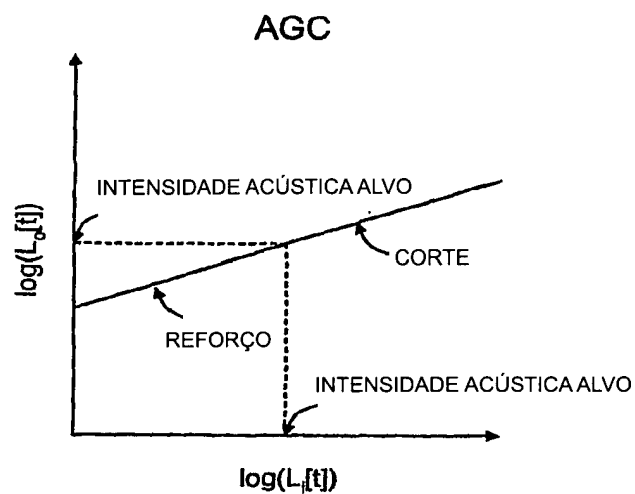


FIG. 14b

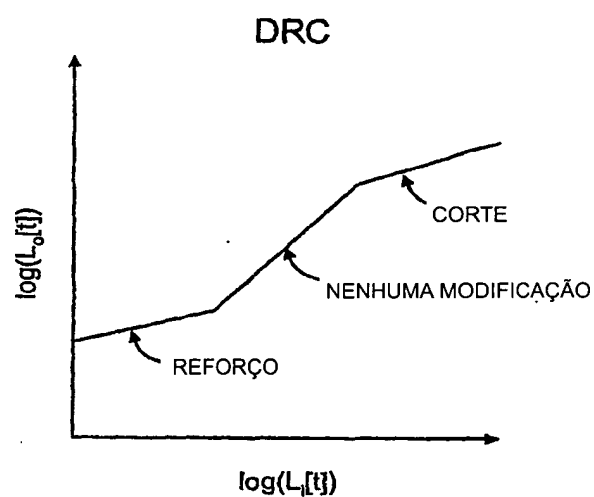
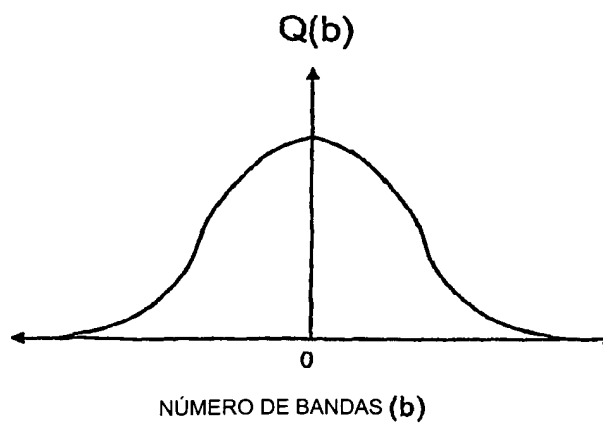


FIG. 15



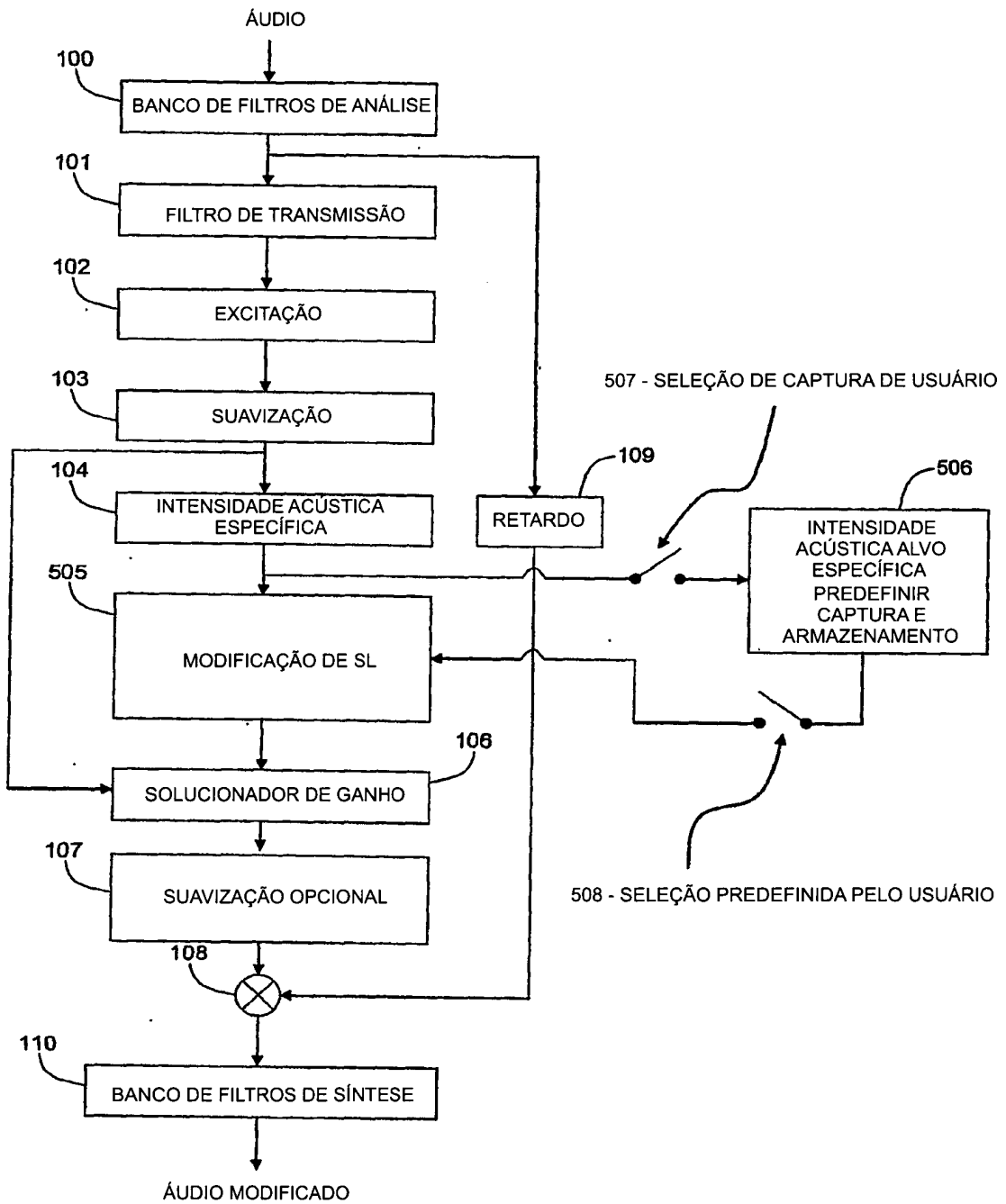


FIG. 16

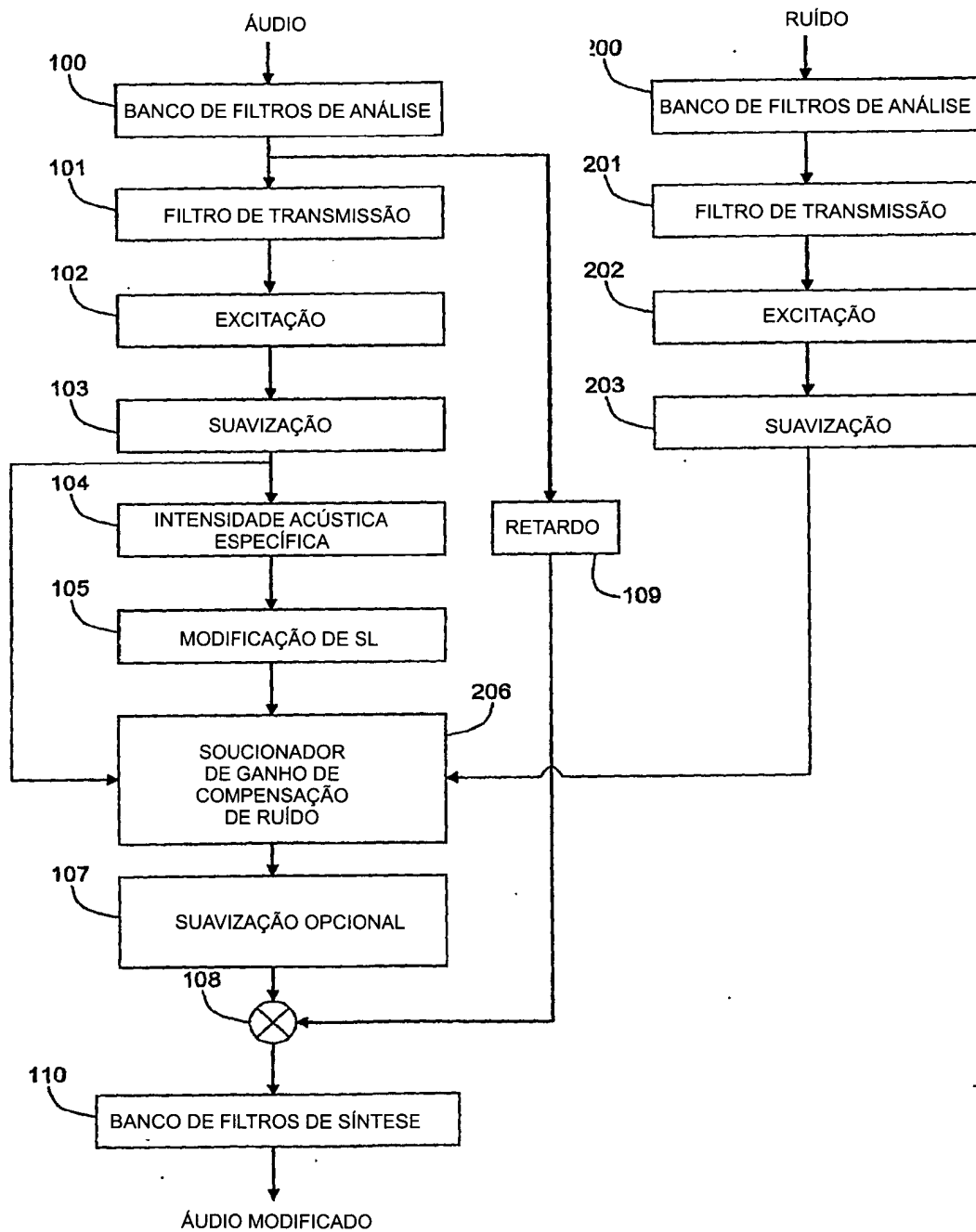


FIG. 17