



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

**(11) BR 112016014949-1 B1**



**(22) Data do Depósito: 18/12/2014**

**(45) Data de Concessão: 22/03/2022**

---

**(54) Título:** MÉTODO E SISTEMA PARA GERAR UM SINAL DE ÁUDIO BINAURAL EM RESPOSTA A ÁUDIO DE MÚLTIPLOS CANAIS AO USAR PELO MENOS UMA REDE DE RETARDO DE FEEDBACK

**(51) Int.Cl.:** H04S 3/00.

**(30) Prioridade Unionista:** 03/01/2014 US 61/923,579; 05/05/2014 US 61/988,617; 29/04/2014 CN 201410178258.0.

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

**(72) Inventor(es):** KUAN-CHIEH YEN; DIRK J. BREEBAART; GRANT A. DAVIDSON; RHONDA WILSON; DAVID M. COOPER; ZHIWEI SHUANG.

**(86) Pedido PCT:** PCT US2014071100 de 18/12/2014

**(87) Publicação PCT:** WO 2015/102920 de 09/07/2015

**(85) Data do Início da Fase Nacional:** 24/06/2016

**(57) Resumo:** GERAÇÃO DE ÁUDIO BINAURAL EM RESPOSTA A ÁUDIO DE MÚLTIPLOS CANAIS AO USAR PELO MENOS UMA REDE DE RETARDO DE FEEDBACK. A presente invenção refere-se, em algumas modalidades, a métodos de virtualização para gerar um sinal binaural em resposta aos canais de um sinal de áudio de múltiplos canais, que aplicam uma res-posta de impulso de ambiente binaural (BRIR) a cada canal incluindo o uso de pelo menos uma rede de retardo de feedback (FDN) para aplicar uma reverberação retardada comum a uma mixagem descendente dos canais. Em algumas modalidades, os canais de sinal da entrada são processados em um primeiro trajeto de processamento para aplicar a cada canal um uma resposta direta e a porção de reflexão antecipada de uma BRIR de um só canal para o canal, e a mixagem descendente dos canais é processada em um segundo trajeto de processamento que inclui pelo menos uma FDN que aplica a reverberação retardada comum. Tipicamente, a reverberação retardada comum emula macroatributos coletivos de poções de reverberação retardada de pelo menos algumas das BRIRs de um só canal. Outros aspectos são virtualizadores de fones auriculares configurados para executar qualquer modalidade do método.

Relatório Descritivo da Patente de Invenção para **"MÉTODO E SISTEMA PARA GERAR UM SINAL DE ÁUDIO BINAURAL EM RESPOSTA A ÁUDIO DE MÚLTIPLOS CANAIS AO USAR PELO MENOS UMA REDE DE RETARDO DE FEEDBACK"**.

Referência Cruzada ao Pedido de Patente Relacionado

[001] O presente pedido de patente reivindica a prioridade do Pedido de Patente Chinês no. 201410178258.0 depositado em 29 de abril de 2014; do Pedido de Patente Provisório U.S. no. 61/923.579 depositado em 03 de janeiro de 2014; e do Pedido de Patente Provisório U.S. no. 61/988.617 depositado em 05 de maio de 2014, cada um dos quais é incorporado no presente documento a título de referência em sua totalidade.

ANTECEDENTES DA INVENÇÃO

1. Campo da Invenção

[002] A invenção refere-se aos métodos (indicados algumas vezes como métodos de virtualização de fones auriculares) e aos sistemas para a geração de um sinal binaural em resposta a um sinal de entrada de áudio de múltiplos canais, mediante a aplicação de uma resposta de impulso de ambiente binaural (BRIR) a cada canal de um conjunto de canais (por exemplo, a todos os canais) do sinal de entrada. Em algumas modalidades, pelo menos uma rede de retardo de feedback (FDN) aplica uma porção de reverberação retardada de uma BRIR de mixagem descendente a uma mixagem descendente dos canais.

2. Antecedentes da Invenção

[003] A virtualização de fones auriculares (ou renderização binaural) é uma tecnologia cujo objetivo é a aplicação de uma experiência com o som circundante ou um campo de som imersivo ao usar fones auriculares estéreo padrão.

[004] Os primeiros virtualizadores de fones auriculares aplicavam uma função de transferência relacionada à cabeça (HRTF) para

conduzir informações espaciais em renderização binaural. Uma HRTF é um conjunto de pares de filtros dependentes da direção e da distância que caracterizam como o som é transmitido de um ponto específico no espaço (localização da fonte do som) a ambas as orelhas de um ouvinte em um ambiente antieco. Indicadores espaciais essenciais tais como a diferença de tempo interaural (ITD), a diferença de nível interaural (ILD), o efeito de sombreamento da cabeça, os picos e entalhes espectrais devidos às reflexões do ombro e da barbatana, podem ser percebidos no conteúdo binaural filtrado por HRTF renderizado. Devido à restrição do tamanho da cabeça humana, as HRTFs não fornecem indicadores suficientes ou robustos a respeito da distância da fonte além de aproximadamente um metro. Em consequência disto, os virtualizadores baseados unicamente em uma HRTF normalmente não obtêm uma boa externalização ou a distância percebida.

[005] A maioria dos eventos acústicos em nossa vida diária acontece nos ambientes reverberantes em que, além do trajeto direto (da fonte à orelha) modelado por HRTF, os sinais de áudio também atingem as orelhas de um ouvinte através de vários trajetos de reflexão. As reflexões introduzem um impacto profundo à percepção auditiva, tal como a distância, o tamanho do ambiente e os outros atributos do espaço. Para conduzir essa informação em renderização binaural, um virtualizador precisa aplicar a reverberação do ambiente além dos indicadores na HRTF de trajeto direto. Uma resposta de impulso de ambiente binaural (BRIR) caracteriza a transformação de sinais de áudio de um ponto específico no espaço às orelhas do ouvinte em um ambiente acústico específico. Em teoria, as BRIRs incluem todos os indicadores acústicos no que diz respeito à percepção espacial.

[006] A figura 1 é um diagrama de blocos de um tipo de virtualizador convencional de fones auriculares que é configurado para aplicar uma resposta de impulso de ambiente binaural (BRIR) a cada

canal de faixa de frequência total ( $X_1, \dots, X_N$ ) de um sinal de entrada de áudio de múltiplos canais. Cada um dos canais  $X_1, \dots, X_N$  é um canal de alto-falante que corresponde a uma direção diferente da fonte em relação a um suposto ouvinte (isto é, a direção de um trajeto direto de uma posição suposta de um alto-falante correspondente à posição do suposto ouvinte), e cada um de tais canais é convolvido pela BRIR para a direção da fonte correspondente. A passagem acústica de cada canal precisa ser simulada para cada orelha. Portanto, no restante do presente documento, o termo BRIR irá se referir a uma resposta de impulso, ou a um par de respostas de impulso associadas com as orelhas esquerda e direita. Desse modo, o subsistema 2 é configurado para convolver o canal  $X_1$  com  $BRIR_1$  (a BRIR para a direção da fonte correspondente), o subsistema 4 é configurado para convolver o canal  $X_N$  com  $BRIR_N$  (a BRIR para a direção da fonte correspondente), e assim por diante. A saída de cada subsistema de BRIR (cada um dos subsistemas 2..., 4) é um sinal do domínio do tempo que inclui um canal esquerdo e um canal direito. As saídas do canal esquerdo dos subsistemas de BRIR são mixadas no elemento de adição 6, e as saídas do canal direito dos subsistemas de BRIR são mixadas no elemento de adição 8. A saída do elemento 6 é o canal esquerdo, L, do sinal de áudio binaural oriundo do virtualizador, e a saída do elemento 8 é o canal direito, R, do sinal áudio binaural oriundo do virtualizador.

[007] O sinal de entrada de áudio de múltiplos canais também pode incluir um efeito de baixa frequência (LFE) ou canal de subwoofer, identificado na figura 1 como o canal de "LFE". De uma maneira convencional, o canal de LFE não é convolvido com uma BRIR, mas é atenuado de preferência no estágio de ganho 5 da figura 1 (por exemplo, por -3dB ou mais) e a saída do estágio de ganho 5 é mixada igualmente (pelos elementos 6 e 8) em cada uma dos canais de sinal de saída binaural do virtualizador. Um estágio de retardo adicional pode ser

necessário no trajeto de LFE a fim de alinhar temporalmente a saída do estágio 5 com as saídas dos subsistemas de BRIR (2,..., 4). Alternativamente, o canal de LFE pode simplesmente ser ignorado (isto é, não reportado a nem processado pelo virtualizador). Por exemplo, a modalidade da figura 2 da invenção (a ser descrito a seguir) ignora simplesmente qualquer canal de LFE do sinal de entrada de áudio de múltiplos canais processado desse modo. Muitos fones auriculares do consumidor não são capazes de reproduzir com exatidão um canal de LFE.

[008] Em alguns virtualizadores convencionais, o sinal de entrada é submetido a uma transformação do domínio do tempo para o domínio da frequência no domínio de QMF (filtro de espelho de quadratura), para gerar os canais de componentes da frequência do domínio de QMF. Esses componentes da frequência são submetidos a uma filtragem (por exemplo, em implementações do domínio de QMF dos subsistemas 2,..., 4 da figura 1) no domínio de QMF e os componentes da frequência resultantes são então transformados tipicamente de volta no domínio do tempo (por exemplo, em um estágio final de cada um dos subsistemas 2,..., 4 da figura 1) de modo que a saída de áudio dos virtualizador seja um sinal do domínio do tempo (por exemplo, sinal binaural do domínio do tempo).

[009] De modo geral, cada canal da faixa de frequência total de uma entrada de sinal áudio de múltiplos canais para um virtualizador de fones auriculares é supostamente indicativo do conteúdo de áudio emitido de uma fonte de som em uma localização conhecida em relação às orelhas do ouvinte. O virtualizador de fones auriculares é configurado para aplicar uma resposta de impulso de ambiente binaural (BRIR) a cada um de tais canais de sinal de entrada. Cada BRIR pode ser decomposta em duas porções: resposta direta e reflexões. A resposta direta é a HRTF que corresponde à direção da chegada (DOA) da fonte

de som, ajustada com o ganho e retardo apropriados devido à distância (entre a fonte de som e o ouvinte), e aumentada opcionalmente com efeitos de paralaxe para distâncias pequenas.

[0010] A porção restante da BRIR modela as reflexões. As reflexões antecipadas são geralmente reflexões primárias ou secundárias e têm uma distribuição temporal relativamente escassa. A microestrutura (por exemplo, ITD e ILD) de cada reflexão primária ou secundária é importante. Para as reflexões retardadas (som refletido de mais de duas superfícies antes de ser incidente no ouvinte), a densidade do eco aumenta com o número crescente das reflexões, e os microatributos de reflexões individuais ficam difíceis de observar. Para as reflexões cada vez mais retardadas, a macroestrutura (por exemplo, a taxa de deterioração da reverberação, a coerência interaural e a distribuição espectral da reverberação total) se torna mais importante. Por causa disso, as reflexões também podem ser segmentadas em duas porções: reflexões antecipadas e reverberações retardadas.

[0011] O retardo da resposta direta é a distância da fonte do ouvinte dividida pela velocidade do som, e o seu nível é (na ausência de paredes ou de grandes superfícies perto da localização da fonte) inversamente proporcional à distância da fonte. Por outro lado, o retardo e o nível das reverberações retardadas são em geral insensíveis à localização da fonte. Devido a considerações práticas, os virtualizadores pode optar por alinhar temporalmente as respostas diretas das fontes com distâncias diferentes, e/ou comprimir a sua faixa dinâmica. No entanto, a relação de sincronismo e de nível entre a resposta direta, as reflexões antecipadas e a reverberação retardada dentro de uma BRIR deve ser mantida.

[0012] O comprimento eficaz de uma BRIR típica estende-se a centenas de milissegundos ou mais na maioria dos ambientes acústicos. A aplicação direta de BRIRs requer a convolução com um

filtro de milhares de derivações, o que é computacionalmente caro. Além disso, sem parameterização, deve requerer um grande espaço de memória para armazenar as BRIRs para uma posição de fonte diferente a fim de obter a definição espacial suficiente. Por último, porém sem menor importância, as localizações da fonte de som podem mudar com o passar do tempo, e/ou a posição e a orientação do ouvinte podem variar com o passar do tempo. A simulação exata de tal movimento requer respostas de impulso de BRIR temporalmente variáveis. A interpolação e a aplicação apropriadas de tais filtros temporalmente variáveis podem ser desafiadoras se as respostas do impulso desses filtros tiverem muitas derivações.

[0013] Um filtro que tem a estrutura de filtro bem conhecida que é conhecida como uma rede de retardo de feedback (FDN) pode ser usado para implementar um reverberator espacial que é configurado para aplicar a reverberação simulada a um ou mais canais de um sinal de entrada de áudio de múltiplos canais. A estrutura de uma FDN é simples. Ela compreende vários tanques de reverberação (por exemplo, o tanque de reverberação que compreende o elemento de ganho  $g_1$  e a linha de retardo  $z^{-n_1}$ , na FDN da figura 4), em que cada tanque de reverberação tem um retardo e um ganho. Em uma implementação típica de uma FDN, as saídas de todos os tanques de reverberação são mixadas por uma matriz de feedback unitária e as saídas da matriz são realimentadas e somadas com as entradas aos tanques de reverberação. Os ajustes de ganho podem ser feitos nas saídas do tanque de reverberação, e as saídas do tanque de reverberação (ou as versões ajustadas em ganho das mesmas) podem ser apropriadamente remixadas para a reprodução de múltiplos canais ou binaural. A reverberação sonora natural pode ser gerada e aplicada por uma FDN com pegadas compactas computacionais e da memória. Portanto, as FDNs são usadas nos virtualizadores para suplementar a resposta

direta produzida pela HRTF.

[0014] Por exemplo, o virtualizador de fone auricular Dolby Mobile comercialmente disponível inclui um reverberador que tem uma estrutura baseada em FDN que é operável para aplicar reverberação a cada canal de um sinal de áudio de cinco canais (que tem canais esquerdo-anterior, direito-anterior, central, esquerdo-circundante e direito-circundante) e para filtrar cada canal reverberado ao usar um par de filtros diferentes de um conjunto de cinco pares de filtros de função de transferência relacionada à cabeça ("HRTF"). O virtualizador de fone auricular Dolby Mobile também é operável em resposta a um sinal de entrada de áudio de dois canais, para gerar uma saída de áudio binaural "reverberada" de dois canais (uma saída de som circundante virtual de dois canais à qual a reverberação foi aplicada). Quando a saída binaural reverberada é renderizada e reproduzida por um par de fones auriculares, ela é percebida nos tímpanos do ouvinte como som reverberado filtrado por HRTF de cinco alto-falantes na parte esquerda anterior, na parte direita anterior, no centro, e na parte esquerda posterior (circundante), e na parte direita posterior (circundante). O virtualizador mixa de modo ascendente uma entrada de áudio de dois canais mixada de modo descendente (sem usar qualquer parâmetro de indicador espacial recebido com a entrada de áudio) para gerar cinco canais de áudio mixados de modo ascendente, aplica a reverberação aos canais mixados de modo ascendente, e mixa de modo descendente os sinais dos cinco canais reverberados para gerar a saída reverberada de dois canais do virtualizador. A reverberação para cada um dos canais mixados de modo ascendente é filtrado em um par de filtros de HRTF diferentes.

[0015] Em um virtualizador, uma FDN pode ser configurada para obter certo tempo de deterioração da reverberação e densidade de eco. No entanto, a FDN não possui a flexibilidade para simular a

microestrutura das reflexões antecipadas. Além disso, nos virtualizadores convencionais a sintonia e a configuração das FDNs são principalmente heurísticas.

[0016] Os virtualizadores de fones auriculares que não simulam todos os trajetos da reflexão (antecipada e retardada) não podem obter uma externalização eficaz. Os autores da presente invenção reconheceram que os virtualizadores que empregam FDNs que tentam simular todos os trajetos de reflexão (antecipada e retardada) têm normalmente um sucesso não mais do que limitado na simulação de reflexões antecipadas e de reverberação retardada e na aplicação de ambas a um sinal de áudio. Os autores da presente invenção também reconheceram que os virtualizadores que empregam FDNs mas não têm a capacidade de controlar corretamente atributos acústicos espaciais tais como o tempo de deterioração da reverberação, a coerência interaural, e a razão direta a retardada, podem atingir um grau de externalização mas à custa de introduzir distorção timbral e reverberação em excesso.

#### BREVE DESCRIÇÃO DA INVENÇÃO

[0017] Em uma primeira classe de modalidades, a invenção é um método para a geração de um sinal binaural em resposta a um conjunto de canais (por exemplo, cada um dos canais, ou cada um dos canais da faixa de frequência total) de um sinal de entrada de áudio de múltiplos canais, o qual inclui as etapas de: (a) aplicação de uma resposta de impulso de ambiente binaural (BRIR) a cada canal do conjunto (por exemplo, a convolução de cada canal do conjunto com uma BRIR que corresponde ao dito canal), desse modo gerando sinais filtrados, inclusive o uso de pelo menos uma rede de retardo de feedback (FDN) para aplicar uma reverberação retardada comum a uma mixagem descendente (por exemplo, uma mixagem descendente monofônica) dos canais do conjunto; e (b) combinação dos sinais filtrados para gerar

o sinal binaural. Tipicamente, um banco de FDNs é usado para aplicar a reverberação retardada comum à mixagem descendente (por exemplo, em que cada FDN aplica a reverberação retardada comum a uma faixa de frequência diferente). Tipicamente, a etapa (a) inclui uma etapa de aplicação, a cada canal do conjunto, de uma porção de "resposta direta e reflexão antecipada" de uma BRIR de um só canal para o canal, e a reverberação retardada comum foi gerada para emular atributos macrocoletivos de porções de reverberação retardada de pelo menos algumas (por exemplo, todas) BRIRs de um só canal.

[0018] Um método para a geração de um sinal binaural em resposta a um sinal de entrada de áudio de múltiplos canais (ou em resposta a um conjunto de canais de tal sinal) é indicado algumas vezes no presente documento como um método "de virtualização de fones auriculares", e um sistema configurado para executar tal método é indicado algumas vezes no presente documento como um "virtualizador de fones auriculares" (ou "sistema de virtualização de fones auriculares" ou "virtualizador binaural").

[0019] Em modalidades típicas na primeira classe, cada uma das FDNs é implementada em um domínio de banco de filtros (por exemplo, o domínio do filtro de espelho de quadratura complexo híbrido (HCQMF) ou o domínio de filtro de espelho da quadratura (QMF), ou outro domínio de transformação ou de sub-banda que pode incluir a decimação) e, em algumas de tais modalidades, os atributos acústicos espaciais dependentes da frequência do sinal binaural são controlados ao controlar a configuração de cada FDN empregada para aplicar a reverberação retardada. Tipicamente, uma mixagem descendente monofônica dos canais é usada como entrada para as FDNs para a renderização binaural eficiente do conteúdo de áudio do sinal de múltiplos canais. As modalidades típicas na primeira classe incluem uma etapa de ajuste dos coeficientes de FDN que correspondem aos

atributos dependentes da frequência (por exemplo, tempo de deterioração da reverberação, coerência interaural, densidade modal, e razão direta a retardada), por exemplo, ao aplicar valores de controle à rede de retardo de feedback para ajustar pelo menos um dentre o ganho de entrada, ganhos do tanque de reverberação, retardos do tanque de reverberação, ou parâmetros de matriz de saída para cada FDN. Isso permite uma melhor combinação de ambientes acústicos e saídas sonoras mais naturais.

[0020] Em uma segunda classe das modalidades, a invenção é um método para a geração de um sinal binaural em resposta a um sinal de entrada de áudio de múltiplos canais que tem canais, mediante a aplicação de uma resposta de impulso de ambiente binaural (BRIR) a cada canal de um conjunto de canais do sinal de entrada (por exemplo, cada um dos canais de sinal de entrada ou de cada canal da faixa de frequência total do sinal de entrada), o qual inclui: o processamento de cada canal do conjunto em um primeiro trajeto de processamento configurado para modelar, e aplicar ao dito cada canal, uma resposta direta e a porção antecipada da reflexão de uma BRIR de um só canal para o canal; e o processamento de uma mixagem descendente (por exemplo, uma (mono) mixagem descendente monofônica) dos canais do conjunto em um segundo trajeto de processamento (em paralelo com o primeiro trajeto de processamento) configurado para modelar, e aplica uma reverberação retardada comum à mixagem descendente. Tipicamente, a reverberação retardada comum foi gerada para emular macroatributos coletivos de porções de reverberação retardada de pelo menos algumas (por exemplo, todas) BRIRs de um só canal. Tipicamente, o segundo trajeto de processamento inclui pelo menos uma FDN (por exemplo, uma FDN para cada uma das faixas de frequência múltiplas). Tipicamente, uma monomixagem descendente é usada como entrada para todos os tanques de reverberação de cada

FDN implementada pelo segundo trajeto de processamento. Tipicamente, são providos mecanismos para o controle sistemático de macroatributos de cada FDN a fim de simular melhores ambientes acústicos e produzir uma virtualização binaural sonora mais natural. Uma vez que a maioria de tais macroatributos é dependente da frequência, cada FDN é tipicamente implementada no domínio do filtro de espelho de quadratura complexo híbrido (HCQMF), no domínio da frequência, no domínio, ou em um outro domínio do banco de filtros, e uma FDN diferente ou independente é usada para cada faixa de frequência. Um benefício primário da implementação das FDNs em um domínio do banco de filtros consiste em permitir a aplicação de reverberação com propriedades de reverberação dependentes da frequência. Em várias modalidades, as FDNs são implementadas em qualquer um de uma ampla variedade de domínios do banco de filtros, ao usar qualquer um de uma variedade de banco de filtros, incluindo, mas sem ficar a eles limitados, filtros de espelho de quadratura (QMF) avaliados reais ou complexos, filtros de resposta de impulso finito (filtros FIR), filtros de resposta de impulso infinito (filtros IIR), transformações de Fourier distintas (DFTs), transformações de cosseno ou seno (modificadas), transformações de Wavelet, ou filtros de cruzamento. Em uma implementação preferida, o banco de filtros empregado ou a transformação inclui a decimação (por exemplo, uma diminuição da taxa da amostragem da representação do sinal do domínio da frequência) para reduzir a complexidade computacional do processo da FDN.

[0021] Algumas modalidades na primeira classe (e na segunda classe) implementam uma ou mais das seguintes características:

[0022] 1. uma implementação de FDN do domínio do banco de filtros (por exemplo, domínio do filtro de espelho de quadratura complexo híbrido), ou a implementação de FDN do domínio de banco de filtros híbrido e a implementação de filtro de reverberação retardada

do domínio do tempo, que permite tipicamente o ajuste independente dos parâmetros e/ou ajustes da FDN para cada faixa de frequência (o que permite o controle simples e flexível de atributos acústicos dependentes da frequência), por exemplo, ao prover a capacidade de variar os retardos do tanque de reverberação em faixas diferentes de modo a mudar a densidade modal como uma função da frequência;

[0023] 2. o processo de mixagem descendente específico, empregado para gerar (a partir do sinal de áudio de entrada de múltiplos canais) o sinal mixado de modo descendente (por exemplo, mixado de modo descendente monofônica) processado no segundo trajeto de processamento, depende da distância da fonte de cada canal e da manipulação da resposta direta a fim de manter o nível apropriado e a relação de sincronismo entre as respostas direta e retardada;

[0024] 3. um filtro de passagem de alta e baixa (APF) é aplicado no segundo trajeto de processamento (por exemplo, na entrada ou na saída de um banco de FDNs) para introduzir a diversidade de fase e a densidade de eco aumentada sem mudar o espectro e/ou o timbre de reverberação resultante;

[0025] 4. retardos fracionários são implementados no trajeto de feedback de cada FDN em uma estrutura de múltiplas taxas avaliada complexa para superar os problemas relacionados com os retardos quantificados para a grade do fator de amostra descendente;

[0026] 5. nas FDNs, as saídas do tanque de reverberação são linearmente mixadas diretamente nos canais binaurais, ao usar os coeficientes de mixagem de saída que são ajustados com base na coerência interaural desejada em cada faixa de frequência. Opcionalmente, o mapeamento dos tanques de reverberação para os canais de saída binaurais é alternante através das faixas de frequência para obter um retardo equilibrado entre os canais binaurais. Também opcionalmente, fatores de normalização são aplicados às saídas do

tanque de reverberação para equalizar os seus níveis enquanto são conservados o retardo fracionário e a potência total;

[0027] 6. o tempo de deterioração da reverberação dependente da frequência e/ou densidade modal é controlado mediante o ajuste das combinações apropriadas dos retardos do tanque de reverberação e ganhos em cada faixa de frequência para simular ambientes reais;

[0028] 7. um fator de escalonamento é aplicado por faixa de frequência (por exemplo, na entrada ou na saída do trajeto de processamento relevante), para:

[0029] controlar uma razão direta a retardada (DLR) dependente da frequência que combina com aquela de um ambiente real (um modelo simples pode ser usado para computar o fator de escalonamento requerido com base na DLR alvo e no tempo de deterioração da reverberação, por exemplo, T60);

[0030] prover a atenuação de baixa frequência para mitigar o excesso de artefatos de coleta e/ou estrondo de baixa frequência; e/ou

[0031] aplicar a moldagem espectral de campo difuso às respostas de FDN;

[0032] 8. modelos paramétricos simples são implementados para controlar os atributos dependentes da frequência essenciais da reverberação retardada, tais como o tempo de deterioração da reverberação, a coerência interaural, e/ou a razão direta a retardada.

[0033] Os aspectos da invenção incluem os métodos e os sistemas que executam (ou são configurados para executar, ou suportar o desempenho de) a virtualização binaural de sinais de áudio (por exemplo, os sinais de áudio cujo conteúdo de áudio consiste em canais de alto-falante, e/ou sinais de áudio baseados em objetos).

[0034] Em uma outra classe de modalidades, a invenção é um método e um sistema para a geração de um sinal binaural em resposta a um conjunto de canais de um sinal de entrada de áudio de múltiplos

canais, o qual inclui a aplicação de uma resposta de impulso de ambiente binaural (BRIR) a cada canal do conjunto, desse modo gerando sinais filtrados, incluindo o uso de uma única rede de retardo de feedback (FDN) para aplicar uma reverberação retardada comum a uma mixagem descendente dos canais do conjunto; e a combinação dos sinais filtrados para gerar o sinal binaural. A FDN é implementada no domínio do tempo. Em algumas de tais modalidades, a FDN do domínio do tempo inclui:

[0035] um filtro de entrada que tem uma entrada acoplada para receber a mixagem descendente, em que o filtro de entrada é configurado para gerar uma primeira mixagem descendente filtrada em resposta à mixagem descendente;

[0036] um filtro de passagem de alta e baixa, acoplado e configurado a uma segunda mixagem descendente filtrada em resposta à primeira mixagem descendente filtrada;

[0037] um subsistema de aplicação de reverberação, que tem uma primeira saída e uma segunda saída, em que o subsistema de aplicação de reverberação compreende um conjunto de tanques de reverberação, em que cada um dos tanques de reverberação tem um retardo diferente, e em que o subsistema de aplicação de reverberação é acoplado e configurado para gerar um primeiro canal binaural não mixado e um segundo canal binaural não mixado em resposta à segunda mixagem descendente filtrada, para aplicar o primeiro canal binaural não mixado na primeira saída, e para aplicar o segundo canal binaural não mixado na segunda saída; e

[0038] um estágio de filtragem e mixagem de coeficiente  $f$  correlação cruzada interaural (IACC) acoplado ao subsistema de aplicação de reverberação e configurado para gerar um primeiro canal binaural mixado e um segundo canal binaural mixado em resposta ao primeiro canal binaural não mixado e um segundo canal binaural não

mixado.

[0039] O filtro de entrada pode ser implementado para gerar (de preferência como uma cascata de dois filtros configurados para gerar) a primeira mixagem descendente filtrada de maneira tal que cada BRIR tem uma razão direta a retardada (DLR) que combina, pelo menos substancialmente, com uma DLR alvo.

[0040] Cada tanque de reverberação pode ser configurado para gerar um sinal retardado, e pode incluir um filtro de reverberação (por exemplo, implementado como um filtro de prateleira ou uma cascata de filtros de prateleira) acoplado e configurado para aplicar um ganho a um sinal que se propaga em cada um dos ditos tanques de reverberação, para fazer com que o sinal retardado tenha um ganho que combine, pelo menos substancialmente, com um ganho de deterioração alvo para o dito sinal retardado, em um esforço para obter uma característica do tempo de deterioração de reverberação alvo (por exemplo, uma característica de T60) de cada BRIR.

[0041] Em algumas modalidades, o primeiro canal binaural não mixado leva ao segundo canal binaural não mixado, os tanques de reverberação incluem um primeiro tanque de reverberação configurado para gerar um primeiro sinal retardado que tem um retardo mais curto e um segundo tanque de reverberação configurado para gerar um segundo sinal retardada que tem um segundo retardo mais curto, em que o primeiro tanque de reverberação é configurado para aplicar um primeiro ganho ao primeiro sinal retardado, o segundo tanque de reverberação é configurado para aplicar um segundo ganho ao segundo sinal retardado, o segundo ganho é diferente do primeiro ganho, o segundo ganho é diferente do primeiro ganho, e a aplicação do primeiro ganho e do segundo ganho resulta na atenuação do primeiro canal binaural não mixado em relação ao segundo canal binaural não mixado. Tipicamente, o primeiro canal binaural mixado e o segundo canal

binaural mixado são indicativos de uma imagem estéreo recentrada. Em algumas modalidades, o estágio de filtragem e mixagem IACC é configurado para gerar o primeiro canal binaural mixado e o segundo canal binaural mixado de maneira tal que o dito primeiro canal binaural mixado e o dito segundo canal binaural mixado têm uma característica de IACC que combina pelo menos substancialmente com uma característica de IACC alvo.

[0042] As modalidades típicas da invenção provêm uma estrutura simples e unificada para suportar ambos o áudio de entrada que consiste em canais do alto-falante, e um áudio de entrada baseado em objetos. Nas modalidades em que as BRIRs são aplicadas aos canais de sinal de entrada que são canais de objetos, o processamento "de resposta direta e reflexão antecipada" executado em cada canal de objetos assume uma direção da fonte indicada pelos metadados providos com o conteúdo de áudio do canal de objetos. Nas modalidades em que as BRIRs são aplicadas aos canais de sinal de entrada que são canais de alto-falantes, o processamento "de resposta direta e reflexão antecipada" executado em cada canal de alto-falante assume uma direção da fonte que corresponde ao canal de alto-falante (isto é, a direção de um trajeto direto de uma posição assumida de um alto-falante correspondente à posição assumida do ouvinte). Independente do fato se os canais de entrada são canais de objetos ou de alto-falantes, o processamento "de reverberação retardada" é executado em uma mixagem descendente (por exemplo, uma mixagem descendente monofônica) dos canais de entrada e não assume nenhuma direção da fonte específica para o conteúdo de áudio da mixagem descendente.

[0043] Outros aspectos da invenção são um virtualizador de fones auriculares configurado (por exemplo, programado) para executar qualquer modalidade do método da invenção, um sistema (por exemplo,

estereofônico, de múltiplos canais, ou um outro decodificador) que inclui tal virtualizador, e um meio que pode ser lido por computador (por exemplo, um disco) que armazena o código para implementar qualquer modalidade do método da invenção.

### BREVE DESCRIÇÃO DOS DESENHOS

[0044] A figura 1 é um diagrama de blocos de um sistema de virtualização de fones auriculares convencional

[0045] A figura 2 é um diagrama de blocos de um sistema que inclui uma modalidade do sistema de virtualização de fones auriculares da invenção.

[0046] A figura 3 é um diagrama de blocos de uma outra modalidade do sistema de virtualização de fones auriculares da invenção.

[0047] A figura 4 é um diagrama de blocos de uma FDN de um tipo incluído em uma implementação típica do sistema da figura 3.

[0048] A figura 5 é um gráfico do tempo de deterioração da reverberação ( $T_{60}$ ) em milissegundos como uma função da frequência em Hz, que pode ser obtido por uma modalidade do virtualizador da invenção para o qual o valor de  $T_{60}$  em cada uma de duas frequências específicas ( $f_A$  e  $f_B$ ) é ajustado tal como segue:  $T_{60,A} = 320$  ms a  $f_A = 10$  Hz, e  $T_{60,b} = 150$  ms a  $f_B = 2,4$  kHz.

[0049] A figura 6 é um gráfico da coerência interaural (Coh) como uma função da frequência em Hz, o qual pode ser obtido por uma modalidade do virtualizador da invenção para o qual os parâmetros de controle  $Coh_{max}$ ,  $Coh_{min}$ , e  $f_c$  são ajustados para ter os seguintes valores:  $Coh_{max} = 0,95$ ,  $Coh_{min} = 0,05$ , e  $f_c = 700$  Hz.

[0050] A figura 7 é um gráfico da razão direta a retardada (DLR) com distância da fonte de um metro, em dB, como uma função da frequência em Hz, que pode ser obtido por uma modalidade do virtualizador da invenção para o qual os parâmetros de controle  $DLR_{1K}$ ,  $DLR_{slope}$ ,  $DLR_{min}$ ,  $HPF_{slope}$ , e  $f_T$  são ajustados para ter os seguintes

valores:  $DLR_{1K} = 18$  dB,  $DLR_{slope} = 6$  dB/10x frequência,  $DLR_{min} = 18$  dB,  $HPF_{slope} = 6$  dB/10x frequência, e  $f_T = 200$  Hz.

[0051] A figura 8 é um diagrama de blocos de uma outra modalidade de um subsistema de processamento da reverberação retardada do sistema de virtualização de fones auriculares da invenção.

[0052] A figura 9 é um diagrama de blocos de uma implementação do domínio do tempo de uma FDN, de um tipo incluído em algumas modalidades do sistema da invenção.

[0053] A figura 9A é um diagrama de blocos de um exemplo de uma implementação do filtro 400 da figura 9.

[0054] A figura 9B é um diagrama de blocos de um exemplo de uma implementação do filtro 406 da figura 9.

[0055] A figura 10 é um diagrama de blocos de uma modalidade do sistema de virtualização de fones auriculares da invenção, em que o subsistema de processamento de reverberação retardada 221 é implementado no domínio do tempo.

[0056] A figura 11 é um diagrama de blocos de uma modalidade dos elementos 422, 423 e 424 da FDN da figura 9.

[0057] A figura 11A é um gráfico da resposta de frequência (R1) de uma implementação típica do filtro 500 da figura 11, da resposta de frequência (R2) de uma implementação típica do filtro 501 da figura 11, e da resposta dos filtros 500 e 501 conectados em paralelo.

[0058] A figura 12 é um gráfico de um exemplo de uma característica de IACC (curva "I") que pode ser obtida por uma implementação da FDN da figura 9, e de uma característica de IACC alvo (curva "I<sub>T</sub>").

[0059] A figura 13 é um gráfico de uma característica de T60 que pode ser obtida por uma implementação da FDN da figura 9, ao implementar apropriadamente cada um dos filtros 406, 407, 408 e 409 implementados como um filtro de prateleira.

[0060] A figura 14 é um gráfico de uma característica de T60 que pode ser obtida por uma implementação da FDN da figura 9, ao implementar apropriadamente cada um dos filtros 406, 407, 408 e 409 implementados como uma cascata de dois filtros IIR de prateleira.

#### Notação e Nomenclatura

[0061] Por toda esta descrição, inclusive nas concretizações, a expressão execução de uma operação "em" um sinal ou dados (por exemplo, filtragem, escalonamento, transformação ou aplicação de ganho ao sinal ou aos dados) é usada em um amplo sentido para denotar a execução da operação diretamente no sinal ou nos dados, ou em uma versão processada do sinal ou dos dados (por exemplo, em uma versão do sinal que foi submetido a filtragem preliminar ou a pré-processamento antes do desempenho da operação nos mesmos).

[0062] Por toda esta descrição inclusive nas concretizações, a expressão "sistema" é usada em um amplo sentido para denotar um dispositivo, um sistema ou um subsistema. Por exemplo, um subsistema que implementa um virtualizador pode ser indicado como um sistema do virtualizador, e um sistema que inclui tal subsistema (por exemplo, um sistema que gera X sinais de saída em resposta a entradas múltiplas, em que o subsistema gera M das entradas e as outras entradas  $X - M$  são recebidas de uma fonte externa) também pode ser indicado como um sistema do virtualizador (ou virtualizador).

[0063] Por toda esta descrição inclusive nas concretizações, o termo "processador" é usado em um amplo sentido para denotar um sistema ou um dispositivo programável ou então configurável (por exemplo, com software ou firmware) para executar operações em dados (por exemplo, áudio, ou vídeo ou outros dados de imagem). Os exemplos dos processadores incluem uma disposição de porta programável no campo (ou um outro circuito integrado ou conjunto de chips configurável), um processador de sinal digital programado e/ou

então configurado para executar o processamento encadeado em áudio ou outros dados sonoros, um processador ou computador de finalidades gerais programável, e um microchip ou conjunto de chips do microprocessador programável.

[0064] Por toda esta descrição inclusive nas concretizações, a expressão "banco de filtros de análise" é usada em um amplo sentido para denotar um sistema (por exemplo, um subsistema) configurado para aplicar uma transformação (por exemplo, uma transformação de domínio do tempo em domínio da frequência) em um sinal de domínio do tempo para gerar valores (por exemplo, componentes da frequência) indicativos do conteúdo do sinal de domínio do tempo, em cada frequência de um conjunto de faixas de frequência. Por toda esta descrição inclusive nas concretizações, a expressão "domínio de banco de filtros" é usada em um amplo sentido para denotar o domínio dos componentes da frequência gerados por uma transformação ou por um banco de filtros de análise (por exemplo, o domínio em que tais componentes da frequência são processados). Os exemplos de domínios de banco de filtros incluem (mas sem ficar a eles limitados) o domínio da frequência, o domínio do filtro de espelho de quadratura (QMF), e o domínio do filtro de espelho de quadratura complexo híbrido (HCQMF). Os exemplos da transformação que pode ser aplicada por um banco de filtros de análise incluem (mas sem ficar a eles limitados) uma transformação de cosseno distinta (DCT), uma transformação de cosseno distinta modificada (MDCT), uma transformação de Fourier distinta (DFT), e uma transformação de Wavelet. Os exemplos de banco de filtros de análise incluem (mas sem ficar a eles limitados) filtros de espelho de quadratura (QMF), filtros de resposta de impulso finito (filtros FIR), filtros de resposta de impulso infinito (filtros IIR), filtros de cruzamento, e filtros que têm outras estruturas de taxas múltiplas apropriadas.

[0065] Por toda esta descrição inclusive nas concretizações, o termo "metadado" refere-se a dados separados e diferentes dos dados de áudio correspondentes (conteúdo de áudio de uma corrente de bits que também inclui metadados). Os metadados são associados com os dados de áudio, e indicam pelo menos uma peculiaridade ou característica dos dados de áudio (por exemplo, que tipo(s) de processamento já foi(foram) executado(s), ou deve(m) ser executado(s), nos dados de áudio, ou na trajetória de um objeto indicado pelos dados de áudio). A associação dos metadados com os dados de áudio é síncrona no tempo. Desse modo, os presentes metadados (recebidos mais recentemente ou atualizados) podem indicar que os dados de áudio correspondentes têm contemporaneamente uma característica indicada e/ou compreendem os resultados de um tipo indicado de processamento de dados de áudio.

[0066] Por toda esta descrição inclusive nas concretizações, o termo "acopla" ou "acoplado" é usado para se referir a uma conexão tanto direta quanto indireta. Desse modo, se um primeiro dispositivo emparelhar com um segundo dispositivo, essa conexão pode ser através de uma conexão direta, ou através de uma conexão indireta por meio de outros dispositivos e conexões.

[0067] Por toda esta descrição inclusive nas concretizações, as expressões a seguir têm as seguintes definições:

[0068] falante e alto-falante são usados como sinônimos para denotar qualquer transdutor emissor de som. Esta definição inclui os alto-falantes implementados como múltiplos transdutores (por exemplo, woofer e tweeter);

[0069] alimentação do alto-falante: um sinal de áudio a ser aplicado diretamente a um alto-falante, ou um sinal de áudio que deve ser aplicado a um amplificador e a um alto-falante em série;

[0070] canal (ou "canal de áudio"): um sinal de áudio monofônico.

Tal sinal pode ser tipicamente renderizado de uma maneira tal que é equivalente à aplicação do sinal diretamente a um alto-falante em uma posição desejada ou nominal. A posição desejada pode ser estática, tal como é tipicamente o caso com alto-falantes físicos, ou dinâmica;

[0071] programa de áudio: um conjunto de um ou mais canais de áudio (pelo menos um canal do alto-falante e/ou pelo menos um canal do objeto) e também opcionalmente metadados associados (por exemplo, metadados que descrevem uma apresentação de áudio espacial desejada);

[0072] canal do alto-falante (ou "canal de alimentação do alto-falante"): um canal áudio que é associado com um alto-falante nomeado (em uma posição desejada ou nominal), ou com uma zona do alto-falante nomeado dentro de uma configuração definida do alto-falante. Um canal de alto-falante é renderizado de uma maneira tal que é equivalente à aplicação do sinal de áudio diretamente ao alto-falante nomeado (na posição desejada ou nominal) ou a um alto-falante na zona do alto-falante nomeado;

[0073] canal de objeto: um canal de áudio indicativo do som emitido por uma fonte de áudio (indicada algumas vezes como um "objeto de áudio"). Tipicamente, um canal de objeto determina uma descrição de fonte de áudio paramétrica (por exemplo, os metadados indicativos da descrição da fonte de áudio paramétrica são incluídos em ou providos com o canal de objeto). A descrição da fonte pode determinar o som emitido pela fonte (como uma função do tempo), pela posição aparente (por exemplo, coordenadas espaciais tridimensionais) da fonte como uma função do tempo, e opcionalmente pelo menos um parâmetro adicional (por exemplo, tamanho ou largura de fonte aparente) que caracteriza a fonte;

[0074] programa de áudio baseado em objeto: um programa de áudio que compreende um conjunto de um ou mais canais de objetos (e

opcionalmente que também compreende pelo menos um canal de alto-falante) e também opcionalmente metadados associados (por exemplo, metadados indicativos de uma trajetória de um objeto de áudio que emite o som indicado por um canal de objeto, ou então metadados indicativos de uma apresentação de áudio espacial desejada do som indicada por um canal de objeto, ou metadados indicativos de uma identificação de pelo menos um objeto de áudio que é uma fonte de som indicada por um canal de objeto); e

[0075] renderização: o processo de converter um programa de áudio em uma ou mais alimentações de alto-falantes, ou o processo de converter um programa de áudio em um ou mais alimentações de alto-falantes e de converter a(s) alimentação(ões) de alto-falantes em som ao usar um ou mais alto-falantes (neste último caso, a renderização é indicada algumas vezes no presente documento como renderização "pelo(s)" alto-falante(s)). Um canal de áudio pode ser trivialmente ser renderizado ("em" uma posição desejada) mediante a aplicação do sinal diretamente a um alto-falante físico na posição desejada, ou um ou o mais canais de áudio podem ser renderizados ao usar uma de uma variedade de técnicas de virtualização projetadas para serem substancialmente equivalentes (para o ouvinte) a tal renderização trivial. Neste último caso, cada canal de áudio pode ser convertido em uma ou mais alimentações de alto-falantes para ser aplicado ao alto-falante(s) nas posições conhecidas, que são de modo geral diferente da posição desejada, de maneira tal que o som emitido pelo alto-falante(s) em resposta à(s) alimentação(ões) será percebido como se emitido da posição desejada. Os exemplos de tais técnicas de virtualização incluem a renderização binaural através de fones auriculares (por exemplo, ao usar o processamento de fones auriculares Dolby que simula até 7,1 canais de som circundante para o usuário dos fones auriculares) e a síntese de campo de ondas.

[0076] A notação que um sinal áudio de múltiplos canais é "x.y" ou um sinal de canal "x.y.z" no presente documento denota que o sinal tem "x" canais de alto-falante de frequência total (que correspondem aos alto-falantes posicionados nominalmente no plano horizontal das orelhas do suposto ouvinte), "y" canais de LFE (ou subwoofer) e, opcionalmente também "z" canais de alto-falante sobre a cabeça de frequência total (que correspondem aos alto-falantes posicionados acima da cabeça do suposto ouvinte, por exemplo, em ou perto do teto de um ambiente).

[0077] A expressão "IACC" denota no presente documento o coeficiente interaural em seu sentido usual, que é uma medida da diferença entre os tempos de chegada do sinal de áudio nas orelhas de um ouvinte, indicado tipicamente por um número em uma faixa de um primeiro valor que indica que os sinais de chegada são iguais na magnitude e exatamente fora de fase, a um valor intermediário que indica que os sinais de chegada não têm nenhuma similaridade, a um valor máximo que indica os sinais de chegada idênticos que têm as mesmas amplitude e fase.

#### DESCRIÇÃO DETALHADA DAS MODALIDADES PREFERIDAS

[0078] Muitas modalidades da presente invenção são tecnologicamente possíveis. Será aparente aos elementos versados no estado da técnica da presente invenção como implementar as mesmas. As modalidades do sistema e do método da invenção serão descritas com referência às Figuras 2 a 14.

[0079] A figura 2 é um diagrama de blocos de um sistema (20) que inclui uma modalidade do sistema de virtualização de fones auriculares da invenção. O sistema de virtualização do fones auriculares (indicado algumas vezes como um virtualizador) é configurado para aplicar uma resposta de impulso de ambiente binaural (BRIR) a N canais da faixa de frequência total ( $X_1, \dots, X_N$ ) de um sinal de entrada de áudio de múltiplos

canais. Cada uma dos canais  $X_1, \dots, X_N$ , (que podem ser canais de alto-falantes ou canais de objetos) corresponde a uma direção da fonte e a uma distância específica em relação a um suposto ouvinte, e na figura 2 o sistema é configurado ao convolver cada uma de tais canais por uma BRIR para a direção da fonte e a distância correspondentes.

[0080] O sistema 20 pode ser um decodificador que é acoplado para receber um programa de áudio codificado, e que inclui um subsistema (não mostrado na figura 2) acoplado e configurado para decodificar o programa que inclui mediante a recuperação dos N canais da faixa de frequência total ( $X_1, \dots, X_N$ ) do mesmo e prover os mesmos aos elementos 12, ..., 14 e 15 do sistema de virtualização (que compreende os elementos, 12, ..., 14, 15, 16 e 18, acoplados tal como mostrado). O decodificador pode incluir subsistemas adicionais, alguns dos quais executam as funções não relacionadas à função de virtualização executada pelo sistema de virtualização, e alguns dos quais podem executar as funções relacionadas à função de virtualização. Por exemplo, estas últimas funções podem incluir a extração de metadados do programa codificado, e a provisão dos metadados a um subsistema de controle da virtualização que emprega os metadados para controlar elementos do sistema de virtualizador.

[0081] O subsistema 12 (com o subsistema 15) é configurado para convolver o canal  $X_1$  com  $BRIR_1$  (a BRIR para a direção da fonte e a distância correspondentes), o subsistema 14 (com o subsistema 15) é configurado para convolver o canal  $X_N$  com  $BRIR_N$  (a BRIR para a direção da fonte correspondente), e assim por diante, para cada um dos outros N-2 subsistemas de BRIR. A saída de cada um dos subsistemas 12, ..., 14 e 15 é um sinal do domínio do tempo que inclui um canal esquerdo e um canal direito. Os elementos de adição 16 e 18 são acoplados às saídas dos elementos 12, ..., 14 e 15. O elemento de adição 16 é configurado para combinar (mixar) as saídas do canal direito

dos subsistemas de BRIR, e o elemento de adição 18 é configurado para combinar (mixar) as saídas do canal direito dos subsistemas de BRIR. A saída do elemento 16 é o canal esquerdo, L, do sinal de áudio binaural oriundo do virtualizador da figura 2, e a saída do elemento 18 é o canal direito, R, do sinal de áudio binaural oriundo do virtualizador da figura 2.

[0082] As características importantes de modalidade típicas da invenção são aparentes a partir da comparação da modalidade da figura 2 do virtualizador de fone auricular da invenção com o virtualizador de fone auricular convencional da figura 1. Para a finalidade de comparação, é suposto que os sistemas da figura 1 e da figura 2 são configurados de modo que, quando o mesmo sinal de entrada de áudio de múltiplos canais é aplicado a cada um deles, os sistemas aplicam uma  $BRIR_i$  que tem a mesma resposta direta e porção de reflexão antecipada (isto é, a  $EBRIR_i$  relevante da figura 2) para cada canal da faixa de frequência total,  $X_i$ , do sinal de entrada (embora não necessariamente com o mesmo grau de sucesso). Cada  $BRIR_i$  aplicada pelo sistema da figura 1 ou da figura 2 pode ser decomposta em duas porções: uma resposta direta e uma porção de reflexão antecipada (por exemplo, uma dentre as porções  $EBIR_1, \dots, EBRIR_N$  aplicadas pelos subsistemas 12 a 14 da figura 2), e por uma porção de reverberação retardada. A modalidade da figura 2 (e outras modalidades típicas da invenção) supor que as porções de reverberação retardada das BRIRs de um só canal,  $BRIR_i$ , podem ser compartilhadas através das direções da fonte e desse modo todos os canais, e aplicar desse modo a mesma reverberação retardada (isto é, uma reverberação retardada comum) a uma mixagem descendente de todos os canais da faixa de frequência total do sinal de entrada. Essa mixagem descendente pode ser uma (mono) mixagem descendente monofônica de todos os canais da entrada, mas pode ser alternativamente uma mixagem descendente

estereofônica ou de múltiplos canais obtida dos canais de entrada (por exemplo, de um subconjunto de canais de entrada).

[0083] Mais especificamente, o subsistema 12 da figura 2 é configurado para convolver o canal de sinal de entrada  $X_1$  com  $EBRIR_1$  (a resposta direta e a porção de reflexão antecipada de BRIR para a direção da fonte correspondente), o subsistema 14 é configurado para convolver o canal  $X_N$  com  $EBRIR_N$  (a resposta direta e a porção de reflexão antecipada de BRIR para a direção da fonte correspondente), e assim por diante. O subsistema de reverberação retardada 15 da figura 2 é configurado para gerar uma monomixagem descendente de todos os canais da faixa de frequência total do sinal de entrada, e convolver a mixagem descendente com LBRIR (uma reverberação retardada comum para todos os canais que são mixados de modo descendente). A saída de cada subsistema de BRIR do virtualizador da figura 2 (cada um dos subsistemas 12,..., 14 e 15) inclui um canal esquerdo e um canal direito (de um sinal binaural gerado do canal de alto-falante ou mixagem descendente correspondente). As saídas do canal esquerdo dos subsistemas de BRIR são combinadas (mixadas) no elemento de adição 16, e as saídas do canal direito dos subsistemas de BRIR são combinadas (mixadas) no elemento de adição 18.

[0084] O elemento de adição 16 pode ser implementado para somar simplesmente as amostras de canais binaurais esquerdos correspondentes (as saídas dos canais esquerdos dos subsistemas 12,..., 14 e 15) para gerar o canal esquerdo do sinal de saída binaural, supondo que os ajustes de nível e os alinhamentos temporais são implementados nos subsistemas 12,..., 14 e 15 apropriados. Similarmente, o elemento de adição 18 também pode ser implementado para somar simplesmente as amostras de canais binaurais direitos correspondentes (por exemplo, as saídas dos canais direitos dos subsistemas 12,..., 14 e 15) para gerar o canal direito do sinal de saída

binaural, supondo outra vez que os ajustes de nível e os alinhamentos temporais são implementados nos subsistemas 12,..., 14 e 15 apropriados.

[0085] O subsistema 15 da figura 2 pode ser implementado em qualquer uma de uma variedade de maneiras, mas inclui tipicamente pelo menos um rede de retardo de feedback configurada para aplicar a reverberação retardada comum a uma mixagem descendente monofônica dos canais de sinal de entrada aplicada à mesma. Tipicamente, onde cada um dos subsistemas 12,..., 14 aplica uma resposta direta e uma porção de reflexão antecipada ( $EBRIR_i$ ) de uma BRIR de um só canal para o canal ( $X_i$ ) que processa, a reverberação retardada comum foi gerado para emular pelo menos macroatributos coletivos de porções de reverberação retardadas de pelo menos algumas (por exemplo, todas) as BRIRs de um só canal (cujas "porções de resposta direta e de reflexão antecipada" são aplicadas pelos subsistemas 12,..., 14). Por exemplo, uma implementação do subsistema 15 tem a mesma estrutura que o subsistema 200 da figura 3, que inclui um banco de redes de retardo de feedback (203, 204,..., 205) configurado para aplicar uma reverberação retardada comum a uma mixagem descendente monofônica dos canais de sinal de entrada aplicada à mesma.

[0086] Os subsistemas 12,..., 14 da figura 2 podem ser implementados em qualquer uma de uma variedade de maneiras (tanto no domínio do tempo ou quanto um domínio de banco de filtros), em que a implementação preferida para qualquer aplicação específica depende de várias considerações, tais como (por exemplo) o desempenho, a computação e a memória. Em uma implementação exemplificadora, cada um dos subsistemas 12, ..., 14 é configurado para convolver o canal aplicado ao mesmo com um filtro FIR que corresponde às respostas direta e antecipada associadas com o canal, com ganho e

retardo ajustados corretamente de modo que as saídas dos subsistemas 12,..., 14 possam ser combinadas simples e eficientemente com aquelas do subsistema 15.

[0087] A figura 3 é um diagrama de blocos de uma outra modalidade do sistema de virtualização de fones auriculares da invenção. A modalidade da figura 3 é similar àquela da figura 2, em que dois sinais de domínio do tempo (canais esquerdo e direito) são emitidos do subsistema de processamento de resposta direta e reflexão antecipada 100, e dois sinais do domínio do tempo (canais esquerdo e direito) são emitidos do subsistema de processamento de reverberação retardada 200. O elemento de adição 210 é acoplado às saídas dos subsistemas 100 e 200. O elemento 210 é configurado para combinar (mixar) as saídas do canal esquerdo dos subsistemas 100 e 200 para gerar o canal esquerdo, L, do sinal de áudio binaural emitido do virtualizador da figura 3, e para combinar (mixar) as saídas do canal direito dos subsistemas 100 e 200 para gerar o canal direito, R, do sinal áudio binaural emitido do virtualizador da figura 3. O elemento 210 pode ser implementado para simplesmente somar a saída esquerda correspondente das amostras de canais dos subsistemas 100 e 200 para gerar o canal esquerdo do sinal de saída binaural, e simplesmente somar as amostras de canais direitos correspondentes oriundos dos subsistemas 100 e 200 para gerar o canal direito do sinal de saída binaural, supondo que os ajustes de nível e os alinhamentos temporais apropriados são executados nos subsistemas 100 e 200.

[0088] No sistema da figura 3, os canais,  $X_i$ , do sinal de entrada de áudio de múltiplos canais são dirigidos para, e submetidos ao processamento em dois trajetos de processamento paralelos: um através do subsistema de processamento de resposta direta e reflexão antecipada 100; o outro através do subsistema de processamento de reverberação retardada 200. O sistema da figura 3 é configurado para

aplicar uma  $BRIR_i$  a cada canal,  $X_i$ . Cada  $BRIR_i$  pode ser decomposta em duas porções: uma resposta direta e uma porção de reflexão antecipada (aplicada pelo subsistema 100), e uma porção de reverberação retardada (aplicada pelo subsistema 200). Em operação, o subsistema processamento de resposta direta e reflexão antecipada 100 gera desse modo a resposta direta e as porções de reflexão antecipada do sinal de áudio binaural que é emitido do virtualizador, e o subsistema de processamento de reverberação retardada ("gerador de reverberação retardada") 200 gera desse modo a porção de reverberação retardada do sinal de áudio binaural que é emitido do virtualizador. As saídas dos subsistemas 100 e 200 são mixadas (pelo subsistema de adição 210) para gerar o sinal de áudio binaural, que é tipicamente aplicado do subsistema 210 a um sistema de renderização (não mostrado) no qual se submete a uma renderização binaural para a reprodução por fones auriculares.

[0089] Tipicamente, quando renderizada e reproduzida por um par de fones auriculares, uma saída de sinal de áudio binaural típica do elemento 210 é percebida nos tímpanos do ouvinte como o som de "N" alto-falantes (onde  $N \geq 2$  e N é tipicamente igual a 2, 5 ou 7) em qualquer uma de uma ampla variedade de posições, incluindo posições na frente, atrás e acima do ouvinte. A reprodução dos sinais de saída gerados na operação do sistema da figura 3 pode propiciar ao ouvinte a experiência do som que vem mais de duas (por exemplo, cinco ou sete) fontes "circundantes". Pelo menos algumas dessas fontes são virtuais.

[0090] O subsistema de processamento de resposta direta e reflexão antecipada 100 pode ser implementado em qualquer uma de uma variedade de maneiras (tanto no domínio do tempo quanto em um domínio do banco de filtros), em que a implementação preferida para qualquer aplicação específica depende de várias considerações, tais como (por exemplo) o desempenho, a computação e a memória. Em

uma implementação exemplificadora, o subsistema 100 é configurado para convolver cada canal aplicado ao mesmo com um filtro FIR que corresponde às respostas diretas e antecipadas associadas com o canal, com ganho e retardo ajustados apropriadamente de modo que as saídas dos subsistemas 100 possam combinar simples e eficientemente (no elemento 210) com aquelas do subsistema 200.

[0091] Tal como mostrado na figura 3, o gerador de reverberação retardada 200 inclui o subsistema de mixagem descendente 201, o banco de filtros de análise 202, um banco de FDNs (FDNs 203, 204, ..., e 205), e o banco de filtros de síntese 207, acoplados tal como mostrado. O subsistema 201 é configurado para mixar de modo descendente os canais do sinal de entrada de múltiplos canais em uma monomixagem descendente e o banco de filtros de análise 202 é configurado para aplicar uma transformação à monomixagem descendente para dividir a monomixagem descendente em "K" faixas de frequência, onde K é um número inteiro positivo. Os valores do domínio do banco de filtros (saída do banco de filtros 202) em cada faixa de frequência diferente são aplicados a uma FDN diferente das FDNs 203, 204..., 205 (há "K" dessas FDNs, cada uma delas acoplada e configurada para aplicar uma porção de reverberação retardada de uma BRIR aos valores do domínio do banco de filtros aplicados à mesma). Os valores do domínio do banco de filtros são de preferência decimados no tempo para reduzir a complexidade computacional das FDNs.

[0092] Em princípio, cada canal da entrada (para o subsistema 100 e o subsistema 201 da figura 3) pode ser processado em sua própria FDN (ou um banco de FDNs) para simular a porção de reverberação retardada de sua BRIR. Apesar do fato que a porção de reverberação retardada das BRIRs associadas com localizações diferentes da fonte de som é tipicamente muito diferente em termos de diferenças do quadrado da raiz média nas respostas do impulso, os seus atributos

estatísticos tais como seu espectro de potência média, a sua estrutura de deterioração de energia, a densidade modal, a densidade de pico e outros similares são frequentemente muito similares. Portanto, a porção de reverberação retardada de um conjunto de BRIRs é tipicamente perceptualmente completamente similar através dos canais e, conseqüentemente, é possível usar uma FDN comum ou um banco de FDNs (por exemplo, as FDNs 203, 204,..., 205) para simular a porção de reverberação retardada de duas ou mais BRIRs. Em modalidades típicas, uma FDN comum (ou banco de FDNs) é empregada, e a entrada para a mesma compreende uma ou mais mixagens descendentes construídas a partir dos canais de entrada. Na implementação exemplificadora da figura 2, a mixagem descendente é uma mixagem descendente monofônica (aplicada na saída do subsistema 201) de todos os canais de entrada.

[0093] Com referência à modalidade da figura 2, cada uma das FDNs 203, 204,..., e 205, é implementada no domínio do banco de filtros, e acoplada e configurada para processar uma faixa de frequência diferente dos valores oriundos do banco de filtros de análise 202, para gerar sinais reverberados esquerdo e direito para cada faixa. Para cada faixa, o sinal reverberado esquerdo é uma seqüência de valores do domínio do banco de filtros, e o sinal reverberado direito é uma outra seqüência de valores do domínio do banco de filtros. O banco de filtros de síntese 207 é acoplado e configurado para aplicar uma transformação de domínio da frequência em domínio do tempo a 2K seqüências de valores do domínio do banco de filtros (por exemplo, componentes da frequência do domínio de QMF) oriundos das FDNs, e para montar os valores transformados em um sinal do domínio do tempo do canal esquerdo (indicativo do conteúdo de áudio da monomixagem descendente à qual a reverberação retardada foi aplicada) e um sinal de domínio do tempo do canal direito (também indicativo do conteúdo

de áudio da monomixagem descendente à qual a reverberação retardada foi aplicada). Esses sinais do canal esquerdo e do canal direito são emitidos ao elemento 210.

[0094] Em uma implementação típica, cada uma das FDNs 203, 204,..., e 205, é implementada no domínio de QMF, e o banco de filtros 202 transforma a monomixagem descendente do subsistema 201 no domínio de QMF (por exemplo, o domínio do filtro de espelho de quadratura híbrido complexo (HCQMF)), de modo que o sinal aplicado do banco de filtros 202 a uma entrada de cada uma das FDNs 203, 204,..., e 205 seja uma sequência de componentes da frequência do domínio de QMF. Em tal implementação, o sinal aplicado do banco de filtros 202 a FDN 203 é uma sequência de componentes da frequência do domínio de QMF em uma primeira faixa de frequência, o sinal aplicado do banco de filtros 202 a FDN 204 é uma sequência de componentes da frequência do domínio de QMF em uma segunda faixa de frequência, e o sinal aplicado do banco de filtros 202 a FDN 205 é uma sequência de componentes da frequência do domínio de QMF em uma "K"<sup>a</sup> faixa de frequência. Quando o banco de filtros de análise 202 é implementado desse modo, o banco de filtros de síntese 207 é configurado para aplicar uma transformação de domínio de QMF em domínio do tempo a 2K sequências de componentes da frequência do domínio de QMF de saída das FDNs, para gerar os sinais de domínio do tempo reverberados retardados do canal esquerdo e do canal direito que são emitidos ao elemento 210.

[0095] Por exemplo, se  $K = 3$  no sistema da figura, então há seis entradas para o banco de filtros de síntese 207 (canais esquerdo e direito, que compreendem amostras do domínio da frequência ou do domínio de QMF, oriundos de cada uma das FDNs 203, 204 e 205) e duas saídas de 207 (canais esquerdo e direito, cada um dos quais consiste em amostras do domínio do tempo). Neste exemplo, o banco

de filtros 207 deve ser tipicamente implementado como dois bancos de filtros de síntese: um (ao qual os três canais esquerdos das FDNs 203, 204 e 205 devem ser aplicados) configurado para gerar o sinal do canal esquerdo do domínio do tempo oriundo do banco de filtros 207; e um segundo (ao qual os três canais direitos das FDNs 203, 204 e 205 devem ser aplicados) configurado para gerar o sinal do canal direito do domínio do tempo oriundo do banco de filtros 207.

[0096] Opcionalmente, o subsistema de controle 209 é acoplado a cada uma das FDNs 203, 204,..., 205, e configurado para aplicar parâmetros de controle a cada uma das FDNs para determinar a porção de reverberação retardada (LBRIR) que é aplicada pelo subsistema 200. Os exemplos de tais parâmetros de controle são descritos a seguir. Contempla-se que em algumas implementações o subsistema de controle 209 é operável em tempo real (por exemplo, em resposta aos comandos do usuário aplicados ao mesmo por um dispositivo de entrada) para implementar a variação em tempo real da porção de reverberação retardada (LBRIR) aplicada pelo subsistema 200 à mixagem descendente monofônica dos canais de entrada.

[0097] Por exemplo, se o sinal de entrada para o sistema da figura 2 for um sinal de 5,1 canais (cujos canais da faixa de frequência total estão na seguinte ordem de canais: L, R, C, Ls, Rs), todos os canais da faixa de frequência total têm a mesma distância da fonte, e o subsistema de mixagem descendente 201 pode ser implementado como a matriz de mixagem descendente a seguir, que simplesmente soma os canais da faixa de frequência total para formar uma monomixagem descendente:

$$D = [ 1 \ 1 \ 1 \ 1 \ 1 ]$$

[0098] Após a filtragem de passagem de alta e baixa (no elemento 301 em cada uma das FDNs 203, 204,..., e 205), a monomixagem descendente acima é mixada de modo ascendente aos quatro tanques de reverberação de uma maneira conservadora de energia:

$$U = \begin{bmatrix} 1/\sqrt{4} \\ 1/\sqrt{4} \\ 1/\sqrt{4} \\ 1/\sqrt{4} \end{bmatrix}$$

[0099] Alternativamente (como um exemplo), é possível optar por girar os canais do lado esquerdo para os dois primeiros tanques de reverberação, os canais do lado direito para os dois últimos tanques de reverberação, e o canal central para todos os tanques de reverberação. Neste caso, o subsistema de mixagem descendente 201 deve ser implementado para formar dois sinais de mixagem descendente:

$$D = \begin{bmatrix} 1 & 0 & 1/\sqrt{2} & 1 & 0 \\ 0 & 1 & 1/\sqrt{2} & 0 & 1 \end{bmatrix}$$

[00100] Neste exemplo, a mixagem ascendente para os tanques de reverberação (em cada uma das FDNs 203, 204,..., e 205) é:

$$U = \begin{bmatrix} 1/\sqrt{2} & 0 \\ 1/\sqrt{2} & 0 \\ 0 & 1/\sqrt{2} \\ 0 & 1/\sqrt{2} \end{bmatrix}$$

[00101] Devido ao fato que há dois sinais de mixagem descendente, o filtro de passagem de alta e baixa (no elemento 301 em cada uma das FDNs 203, 204,..., e 205) precisa ser aplicado duas vezes. A diversidade deve ser introduzida para as respostas retardadas de (L, Ls), (R, Rs) e C apesar de todos terem os mesmos macroatributos. Quando os canais de sinal de entrada têm distâncias de fontes diferentes, retardos e ganhos apropriados ainda precisam ser aplicados no processo de mixagem descendente.

[00102] Serão descritas em seguida considerações para implementações específicas do subsistema de mixagem descendente 201, e os subsistemas 100 e 200 do virtualizador da figura 3.

[00103] O processo de mixagem descendente implementado pelo subsistema 201 depende da distância da fonte (entre a fonte de som e a posição do suposto ouvinte) para que cada canal seja mixado de modo

descendente, e a manipulação da resposta direta. O retardo da resposta direta  $t_d$  é:

$$t_d = d / v_s$$

onde  $d$  é a distância entre a fonte de som e o ouvinte e  $v_s$  é a velocidade do som. Além disso, o ganho da resposta direta é proporcional a  $1/d$ . Se estas regras forem preservadas na manipulação de respostas diretas dos canais com distâncias de fontes diferentes, o subsistema 201 pode implementar uma a mixagem descendente direta de todos os canais porque o retardo e o nível de reverberação retardada é geralmente insensível à localização da fonte.

[00104] Devido às considerações práticas, virtualizadores (por exemplo, o subsistema 100 do virtualizador da figura 3) podem ser implementados para alinhar temporalmente as respostas diretas para os canais de entrada que têm distâncias de fontes diferentes. A fim de preservar o retardo relativo entre a resposta direta e a reverberação retardada para cada canal, um canal com a distância de fonte  $d$  deve ser retardado por  $(d_{max} - d)/v_s$  antes de ser mixado de modo descendente com outros canais. Aqui,  $d_{max}$  denota a distância de fonte possível máxima.

[00105] Virtualizadores (por exemplo, o subsistema 100 do virtualizador da figura 3) também podem ser implementados para comprimir a faixa dinâmica das respostas diretas. Por exemplo, a resposta direta para um canal com distância de fonte  $d$  pode ser escalonada por um fator de  $d^{-\alpha}$ , onde  $0 \leq \alpha \leq 1$ , em vez de  $d^{-1}$ . A fim de preservar a diferença de nível entre a resposta direta e a reverberação retardada, a mixagem descendente do subsistema 201 pode ter que ser implementada para escalonar um canal com distância de fonte  $d$  por um fator de  $d^{1-\alpha}$  antes da mixagem descendente do mesmo com outros canais escalonados.

[00106] A rede de retardo de feedback da figura 4 é uma

implementação exemplificadora da FDN 203 (ou 204 ou 205) da figura 3. Embora o sistema da figura 4 tenha quatro tanques de reverberação (cada um incluindo um estágio de ganho,  $g_i$ , e uma linha de retardo,  $z^{-n_i}$ , acoplados à saída do estágio do ganho), suas variações no sistema (e outras FDNs empregadas nas modalidades do virtualizador da invenção) implementam mais de ou menos de quatro tanques de reverberação.

[00107] A FDN da figura 4 inclui o elemento de ganho de entrada 300, o filtro de passagem de alta e baixa (APF) 301 acoplado à saída do elemento 300, os elementos de adição 302, 303, 304 e 305 acoplados à saída do APF 301, e quatro tanques de reverberação (cada um dos quais compreende um elemento de ganho,  $g_k$  (um dos elementos 306), uma linha de retardo,  $Z^{-Mk}$  (um dos elementos 307) acoplada à mesma, e um elemento de ganho,  $1/g_k$  (um dos elementos 309) acoplado à mesma, onde  $0 \leq k - 1 \leq 3$ ) cada um deles acoplado à saída de um elemento diferente dentre os elementos 302, 303, 304 e 305. A matriz unitária 308 é acoplada às saídas das linhas de retardo 307, e configurada para aplicar uma saída de feedback a uma segunda entrada de cada um dos elementos 302, 303, 304 e 305. As saídas de dois dos elementos de ganho 309 (do primeiro e do segundo tanques de reverberação) são aplicadas às entradas do elemento de adição 310, e a saída do elemento 310 é aplicada a uma entrada da matriz de mixagem de saída 312. As saídas dos outros dois dos elementos de ganho 309 (do terceiro e do quarto tanques de reverberação) são aplicadas às entradas do elemento de adição 311, e a saída do elemento 311 é aplicada à outra entrada da matriz de mixagem de saída 312.

[00108] O elemento 302 é configurado para adicionar a saída da matriz 308 que corresponde à linha de retardo  $z^{-n_i}$  (isto é, aplicar o feedback da saída da linha de retardo  $z^{-n_i}$  através da matriz 308) à

entrada do primeiro tanque de reverberação. O elemento 303 é configurado para adicionar a saída da matriz 308 que corresponde à linha de retardo  $z^{-n^2}$  (isto é, aplicar o feedback da saída da linha de retardo  $z^{-n^2}$  através da matriz 308) à entrada do segundo tanque de reverberação. O elemento 304 é configurado para adicionar a saída da matriz 308 que corresponde à linha de retardo  $z^{-n^3}$  (isto é, aplicar o feedback da saída da linha de retardo  $z^{-n^3}$  através da matriz 308) à entrada do terceiro tanque de reverberação. O elemento 305 é configurado para adicionar a saída da matriz 308 que corresponde à linha de retardo  $z^{-n^4}$  (isto é, aplicar o feedback de saída da linha de retardo  $z^{-n^4}$  através da matriz 308) à entrada do quarto tanque de reverberação.

[00109] O elemento de ganho de entrada 300 da FDN da figura 4 é acoplado para receber uma faixa de frequência do sinal de mixagem descendente monofônica transformado (um sinal do domínio do banco de filtros) que é oriundo do banco de filtros de análise 202 da figura 3. O elemento de ganho de entrada 300 aplica um fator de ganho (escalonamento),  $G_{in}$ , ao sinal do domínio do banco de filtros aplicado ao mesmo. Coletivamente, os fatores de escalonamento  $G_{in}$  (implementados por todas as FDNs 203, 204, ..., 205 da figura 3) para todas as faixas de frequência controlam o formato espectral e o nível de reverberação retardada. O ajuste dos ganhos de entrada,  $G_{in}$ , em todas as FDNs do virtualizador da figura 3 leva frequentemente em conta os seguintes objetivos:

[00110] uma razão direta a retardada (DLR), da BRIR aplicada a cada canal, que combina com ambientes reais;

[00111] atenuação de baixa frequência necessária para mitigar o excesso de artefatos de coleta e/ou estrondo de baixa frequência; e

[00112] combinação do envelope espectral de campo difuso.

[00113] Se for suposto que a resposta direta (aplicada pelo

subsistema 100 da figura 3) provê o ganho unitário em todas as faixas de frequência, uma DLR específica (razão de potência) pode ser obtida ao ajustar  $G_{in}$  para que seja:

$$G_{in} = \sqrt{(\ln(10^6)/(T60 * DLR))},$$

onde T60 é o tempo de deterioração da reverberação definido como o tempo que a reverberação leva para deteriorar por 60 dB (é determinado pelos retardos de reverberação e pelos ganhos de reverberação discutidos a seguir), e "ln" denota a função logarítmica natural.

[00114] O fator de ganho de entrada,  $G_{in}$ , pode ser dependente do conteúdo que está sendo processado. Uma aplicação de tal dependência do conteúdo consiste em assegurar que a energia da mixagem descendente em cada segmento de tempo/frequência é igual à soma das energias dos sinais de canais individuais que estão sendo mixados de modo descendente, independentemente de qualquer correlação que puder existir entre os sinais do canal de entrada. Nesse caso, o fator de ganho de entrada pode ser (ou pode ser multiplicado por) um termo similar ou igual a:

$$\sqrt{\frac{\sum_i \sum_j x_i^2(j)}{\sum_j y^2(j)}}$$

em que  $i$  é um índice sobre todas as amostras de mixagem descendente de uma certa telha ou subfaixa de tempo/frequência,  $y(i)$  são as amostras de mixagem descendente para a telha, e  $x_i(j)$  é o sinal de entrada (para o canal  $X_i$ ) aplicado à entrada do subsistema de mixagem descendente 201.

[00115] Em uma implementação típica do domínio de QMF da FDN da figura 4, o sinal de saída aplicado do filtro de passagem de alta e baixa (APF) 301 às entradas dos tanques de reverberação é uma sequência de componentes da frequência do domínio de QMF. Para gerar uma saída de FDN de som mais natural, o APF 301 é aplicado à

saída do elemento de ganho 300 para introduzir a diversidade de fase e a densidade de eco aumentada. Alternativa, ou adicionalmente, um ou mais filtros de retardo de passagem de alta e baixa podem ser aplicados a: entradas individuais ao subsistema de mixagem descendente 201 (da figura 3) antes de serem mixados de modo descendente no subsistema 201 e processados pela FDN; ou as passagens de alimentação para diante ou de alimentação para trás do tanque de reverberação mostradas na figura 4 (por exemplo, além de ou em substituição às linhas de retardo  $z^{-M_i}$  em cada tanque de reverberação; ou as saídas da FDN (isto é, às saídas da matriz de saída 312).

[00116] Na implementação dos retardos do tanque de reverberação,  $z^{-n_i}$ , os retardos de reverberação  $n_i$  devem ser números mutuamente primos para evitar os modos de reverberação que alinham à mesma frequência. A soma dos retardos atrasa deve ser grande o bastante para prover a densidade modal suficiente a fim de evitar a saída de som artificial. Mas os retardos mais curtos devem ser curtos o bastante para evitar um intervalo de tempo excessivo entre a reverberação retardada e os outros componentes da BRIR.

[00117] Tipicamente, as saídas do tanque de reverberação são garimpadas inicialmente para o canal binaural tanto esquerdo quanto direito. Normalmente, os conjuntos das saídas de tanques de reverberação que são garimpadas a dois canais binaural são iguais no número e mutuamente exclusivos. Também é desejável balancear o sincronismo dos dois canais binaurais. Desse modo, se a saída do tanque de reverberação com o retardo mais curto for para um canal binaural, a saída com o segundo retardo mais curto deve ir para o outro canal.

[00118] Os retardos de tanque de reverberação podem ser diferentes através das faixas de frequência de modo a mudar a densidade modal como uma função da frequência. Normalmente, as faixas de frequência

mais baixas requerem uma densidade modal mais elevada, desse modo os retardos de tanque de reverberação mais longos.

[00119] As amplitudes dos ganhos do tanque de reverberação,  $g_i$ , e os retardos do tanque de reverberação, determinam conjuntamente o tempo de deterioração de reverberação da FDN da figura 4:

$$T_{60} = -3n_i / \log_{10}(|g_i|) / F_{FRM}$$

onde  $F_{FRM}$  é a taxa de quadros do banco de filtros 202 (da figura 3). As fases dos ganhos do tanque de reverberação introduzem retardos fracionários para superar os problemas relacionados aos atrasos do tanque de reverberação que são quantificados à grade do fator de amostragem descendente do banco de filtros.

[00120] A matriz de feedback unitária 308 provê até mesmo a mixagem entre os tanques de reverberação no trajeto de feedback.

[00121] Para equalizar os níveis das saídas do tanque de reverberação, os elementos de ganho 309 aplicam um ganho de normalização,  $1/|g_i|$  à saída de cada tanque de reverberação, para remover o impacto de nível dos ganhos do tanque de reverberação enquanto são preservados os retardos fracionários introduzidos por suas fases.

[00122] A matriz de mixagem de saída 312 (também identificada como matriz  $M_{out}$ ) é uma matriz de  $2 \times 2$  configurada para mixar os canais binaurais não mixados (as saídas dos elementos 310 e 311, respectivamente) da garimpagem inicial para obter os canais binaurais esquerdo e direito de saída (os sinais L e R aplicados na saída da matriz 312) que têm a coerência interaural desejada. Os canais binaurais não mixados ficam perto de não ficar correlacionados após a garimpagem inicial porque não consistem em nenhuma saída comum do tanque de reverberação. Se a coerência interaural desejada for  $Coh$ , onde  $|Coh| \leq 1$ , a matriz de mixagem de saída 312 pode ser definida como:

$$M_{out} = \begin{bmatrix} \cos \beta & \sin \beta \\ \sin \beta & \cos \beta \end{bmatrix}, \text{ onde } \beta = \arcsin(Coh)/2$$

[00123] Devido ao fato que os retardos do tanque de reverberação são diferentes, um dos canais binaurais não mixados deve conduzir ao outro constantemente. Se a combinação dos retardos do tanque de reverberação e do padrão de garimpagem for idêntica através das faixas de frequência, deve resultar em uma polarização de som e imagem. Essa polarização pode ser mitigada se o padrão de garimpagem for alternado através das faixas de frequência de maneira tal que os canais binaurais mixados conduzirem e seguirem uns os outros em faixas de frequência alternadas. Isso pode ser obtido ao implementar a matriz de mixagem de saída 312 de modo a ter a forma tal como indicado no parágrafo precedente em faixas de frequência de números ímpares (isto é, na primeira faixa de frequência (processada pela FDN 203 da figura 3), na terceira faixa de frequência, e assim por diante), e para ter a seguinte forma em faixas de frequência de números pares (isto é, na segunda faixa de frequência (processada pela FDN 204 da figura 3), na quarta faixa de frequência, e assim por diante):

$$M_{out,alt} = \begin{bmatrix} \sin \beta & \cos \beta \\ \cos \beta & \sin \beta \end{bmatrix}$$

onde a definição de  $\beta$  permanece a mesma. Deve ser observado que a matriz 312 pode ser implementada para ser idêntica nas FDNs para todas as faixas de frequência, mas a ordem do canal de suas entradas pode ser comutada para alternar uma das faixas de frequência (por exemplo, a saída do elemento 310 pode ser aplicada à primeira entrada da matriz 312 e a saída do elemento 311 pode ser aplicada à segunda entrada da matriz 312 em faixas de frequência ímpares, e a saída do elemento 311 pode ser aplicada à primeira entrada da matriz 312 e a saída do elemento 310 pode ser aplicada à segunda entrada da matriz

312 em faixas de frequência pares.

[00124] No caso que as faixas de frequência estão (parcialmente) sobrepostas, a largura da faixa de frequência na qual a forma da matriz 312 é alternada pode ser aumentada (por exemplo, poderia ser alternada uma vez para cada duas ou três faixas consecutivas), ou o valor de  $\beta$  nas expressões acima (para a forma da matriz 312) pode ser ajustada para assegurar que a coerência média seja igual ao valor desejado para compensar a sobreposição espectral de faixas de frequência consecutivas.

[00125] Se os atributos acústicos alvo definidos acima  $T_{60}$ , Coh e DLR forem conhecidos para a FDN para cada faixa de frequência específica no virtualizador da invenção, cada uma das FDNs (cada uma das quais pode ter a estrutura mostrada na figura 4) pode ser configurada para obter os atributos alvo. Especificamente, em algumas modalidades os ganhos de entrada ( $G_{in}$ ) e os ganhos e retardos do tanque de reverberação ( $g_i$  e  $n_i$ ) e os parâmetros da matriz de saída  $M_{out}$  para cada FDN podem ser ajustados (por exemplo, pelos valores de controle aplicados aos mesmos pelo subsistema de controle 209 da figura 3) para obter os atributos alvo de acordo com as relações descritas no presente documento. Na prática, o ajuste dos atributos dependentes da frequência por modelos com parâmetros de controle simples é frequentemente suficiente para gerar a reverberação retardada de som natural que combina com ambientes acústicos específicos.

[00126] Será descrito em seguida um exemplo de como um tempo de deterioração de reverberação alvo ( $T_{60}$ ) para a FDN para cada faixa de frequência específica de uma modalidade do virtualizador da invenção pode ser determinado, mediante a determinação do tempo de deterioração de reverberação alvo ( $T_{60}$ ) para cada frequência de um número pequeno de faixas de frequência. O nível da resposta de FDN

deteriora exponencialmente com o passar do tempo.  $T_{60}$  é inversamente proporcional ao fator de deterioração,  $d_f$  (definido como a deterioração de dB em relação a uma unidade de tempo):

$$T_{60} = 60/d_f.$$

[00127] O fator de deterioração,  $d_f$ , depende da frequência e de modo geral aumenta linearmente versus a escala de log-frequência, de modo que o tempo de deterioração de reverberação também é uma função da frequência que diminui em geral à medida que a frequência aumenta. Portanto, se forem determinados (por exemplo, ajustados) os valores de  $T_{60}$  para dois pontos da frequência, a curva de  $T_{60}$  para todas as frequências é determinada. Por exemplo, se os tempos de deterioração de reverberação para os pontos de frequência  $f_A$  e  $f_B$  forem  $T_{60,A}$  e  $T_{60,B}$ , respectivamente, a curva de  $T_{60}$  é definida como:

$$T_{60}(f) = \frac{T_{60,A}T_{60,B} \log(f_B / f_A)}{T_{60,A} \log(f / f_A) - T_{60,B} \log(f / f_B)}$$

[00128] A figura 5 mostra um exemplo de uma curva de  $T_{60}$  que pode ser obtida por uma modalidade do virtualizador da invenção para a qual o valor de  $T_{60}$  em cada uma de duas frequências específicas ( $f_A$  e  $f_B$ ) é ajustado:  $T_{60,A} = 320$  ms a  $f_A = 10$  Hz, e  $T_{60,B} = 150$  ms a  $f_B = 2,4$  kHz.

[00129] Será descrito em seguida um exemplo de como uma coerência Interaural alvo (Coh) para a FDN para cada faixa de frequência específica de uma modalidade do virtualizador da invenção pode ser obtida mediante o ajuste de um número pequeno de parâmetros de controle. A coerência Interaural (Coh) de reverberação retardada segue bastante o padrão de um campo sonoro difuso. Ela pode ser modelada por uma função do seno até uma frequência de cruzamento  $f_c$ , e uma constante acima da frequência de cruzamento. Um modelo simples para a curva de Coh é:

$$Coh(f) = \begin{cases} Coh_{\min} + (Coh_{\max} - Coh_{\min}) \text{sinc}(f / f_c) & , f \leq f_c \\ Coh_{\min} & , f \geq f_c \end{cases}$$

onde os parâmetros  $Coh_{\min}$  e  $Coh_{\max}$  satisfazem  $-1 \leq Coh_{\min} < Coh_{\max} \leq 1$ , e controlam a faixa de Coh. A frequência de cruzamento  $f_c$  ideal depende do tamanho da cabeça do ouvinte. Uma  $f_c$  demasiadamente elevada conduz à imagem de fonte de som internalizada, ao passo que um valor demasiadamente pequeno conduz à imagem da fonte de som dispersa ou dividida. A figura 6 é um exemplo de uma curva de Coh que pode ser obtida por uma modalidade do virtualizador da invenção para a qual os parâmetros de controle  $Coh_{\max}$ ,  $Coh_{\min}$  e  $f_c$  são ajustados para que tenham os seguintes valores:  $Coh_{\max} = 0,95$ ,  $Coh_{\min} = 0,05$  e  $f_c = 700$  Hz.

[00130] Será descrito em seguida um exemplo de como uma razão direta a retardada alvo (DLR) para a FDN para cada faixa de frequência específica de uma modalidade do virtualizador da invenção pode ser obtida mediante o ajuste de um número pequeno de parâmetros de controle. A razão direta a retardada (DLR), em dB, de modo geral aumenta linearmente versus a faixa de log-frequência. Ela pode ser controlada mediante o ajuste de  $DLR_{1K}$  (DLR em dB a 1 kHz) e  $DLR_{\text{slope}}$  (em dB por 10x frequência). No entanto, uma DLR baixa na faixa de frequência mais baixa resulta frequentemente em um artefato de coleta excessivo. A fim de mitigar o artefato de coleta, dois mecanismos modificadores são adicionados para o controle da DLR:

[00131] um piso de DLR mínimo,  $DLR_{\min}$  (em dB); e

[00132] um filtro de passagem de alta frequência definido por uma frequência de transição,  $f_T$  e pela inclinação da curva de atenuação abaixo da mesma,  $HPF_{\text{slope}}$  (em dB por 10x frequência).

[00133] A curva de DLR resultante em dB é definida como:

$$DLR(f) = \max(DLR_{1K} + DLR_{slope} \log_{10}(f/1000) , DLR_{min}) + \min(HPF_{slope} \log_{10}(f/f_T) , 0)$$

[00134] Deve ser observado que a DLR muda com a distância da fonte até mesmo no mesmo ambiente acústico. Portanto,  $DLR_{1K}$  e  $DLR_{min}$  aqui são os valores para uma distância nominal da fonte, tal como 1 metro. A figura 7 é um exemplo de uma curva de DLR para a distância da fonte de 1 metro obtida por uma modalidade do virtualizador da invenção com os parâmetros de controle  $DLR_{1K}$ ,  $DLR_{slope}$ ,  $DLR_{min}$ ,  $HPF_{slope}$  e  $f_T$  ajustados para que tenham os seguintes valores:  $DLR_{1K} = DB\ 18$ ,  $DLR_{slope} = 6\ dB/10x\ frequência$ ,  $DLR_{min} = DB\ 18$ ,  $HPF_{slope} = 6\ dB/10x\ frequência$ , e  $f_T = 200\ Hz$ .

[00135] As variações nas modalidades divulgadas no presente documento têm uma ou mais das seguintes características:

[00136] as FDNs do virtualizador da invenção são implementadas no domínio do tempo, ou têm a implementação híbrida com a captura da resposta de impulso baseada em FDN e a filtragem de sinal baseada em FIR;

[00137] o virtualizador da invenção é implementado para permitir a aplicação da compensação de energia como uma função da frequência durante o desempenho da etapa de mixagem descendente que gera o sinal de entrada mixado de modo descendente para o subsistema de processamento de reverberação retardada; e

[00138] o virtualizador da invenção é implementado para permitir o controle manual ou automático dos atributos de reverberação retardada aplicados em resposta a fatores externos (isto é, em resposta ao ajuste dos parâmetros de controle).

[00139] Para as aplicações em que a latência do sistema é crítica e o retardo causado pelo banco de filtros de análise e síntese é proibitivo, a estrutura da FDN do domínio de banco de filtros de modalidade típicas do virtualizador da invenção pode ser traduzida no domínio do tempo, e

cada estrutura da FDN pode ser implementada no domínio do tempo em uma classe de modalidades do virtualizador. Nas implementações do domínio do tempo, os subsistemas que aplicam o fator de ganho de entrada ( $G_{in}$ ), ganhos do tanque de reverberação ( $g_i$ ) e ganhos de normalização ( $1/|g_i|$ ) são substituídos por filtros com respostas de amplitude similares a fim de permitir controles dependentes da frequência. A matriz de mixagem de saída ( $M_{out}$ ) também é substituída por uma matriz de filtros. Ao contrário dos outros filtros, a resposta de fase dessa matriz de filtros é crítica, uma vez que a conservação de energia e a coerência interaural podem ser afetadas pela resposta de fase. Os retardos do tanque de reverberação em uma implementação do domínio do tempo podem ter que ser ligeiramente variados (de seus valores em uma implementação do domínio do banco de filtros) para evitar o compartilhamento do passo do banco de filtros como um fator comum. Devido a várias restrições, o desempenho de implementações do domínio do tempo das FDNs do virtualizador da invenção pode não combinar precisamente com aquelas de suas implementações do domínio do banco de filtros.

[00140] Com referência à figura 8, será descrita em seguida uma implementação híbrida (domínio do banco de filtros e domínio do tempo) do subsistema de processamento de reverberação retardada da invenção do virtualizador da invenção. Essa implementação híbrida do subsistema de processamento de reverberação retardada da invenção é uma variação no subsistema de processamento de reverberação retardada 200 da figura 4, que implementa a resposta de resposta de impulso baseada em FDN e a filtragem de sinal baseada em FIR.

[00141] A modalidade da figura 8 inclui os elementos 201, 202, 203, 204, 205 e 207 que são idênticos aos elementos identicamente numerados do subsistema 200 da figura 3. A descrição acima desses elementos não será repetida com referência à figura 8. Na modalidade

da figura, o gerador de impulso unitário 211 é acoplado para aplicar um sinal de entrada (um pulso) ao banco de filtros de análise 202. Um filtro LBRIR 202 (mono-entrada, estéreo-saída) implementado como um filtro FIR aplica a porção de reverberação retardada apropriada da BRIR (a LBRIR) à mixagem descendente monofônica oriunda do subsistema 201. Desse modo, os elementos 211, 202, 203, 204, 205 e 207 são uma cadeia lateral de processamento para o filtro LBRIR 208.

[00142] Sempre que o ajuste da porção de reverberação retardada LBRIR tiver que ser modificado, o gerador de impulso 211 é operado para aplicar um impulso unitário ao elemento 202, e a saída resultante do banco de filtros 207 é capturada e aplicada ao filtro 208 (para ajustar o filtro 208 para aplicar a nova LBRIR determinada pela saída do banco de filtros 207). Para acelerar a mudança do lapso de tempo da mudança do ajuste de LBRIR para o tempo de que a nova LBRIR faz efeito, as amostras da nova LBRIR podem começar a substituir a antiga LBRIR à medida que se tornam disponíveis. Para encurtar a latência inerente das FDNs, os zeros iniciais da LBRIR podem ser descartados. Essas opções propiciam flexibilidade e permitem que a implementação híbrida confira uma melhora potencial do desempenho (em relação ao que é conferido por uma implementação do domínio do banco de filtros), a um custo da computação adicionada da filtragem FIR.

[00143] Para as aplicações em que a latência do sistema é crítica, mas o poder de computação tem um menor interesse, o processador de reverberação retardada do domínio do banco de filtros de cadeia lateral (por exemplo, que foi implementado pelos elementos 211, 202, 203, 204..., 205 e 207 da figura 8) pode ser usado para capturar a resposta de impulso FIR eficaz a ser aplicada pelo filtro 208. O filtro FIR 208 pode implementar essa resposta de FIR capturada e aplicar a mesma diretamente à monomixagem descendente dos canais de entrada (durante a virtualização dos canais da entrada).

[00144] Os vários parâmetros de FDN e desse modo os atributos da reverberação retardada resultante podem ser manualmente ser ajustados e subsequentemente fiados em uma modalidade do subsistema de processamento de reverberação retardada da invenção, por exemplo, por meio de um ou mais pré-ajustes que podem ser ajustados (por exemplo, pelo subsistema de controle operacional 209 da figura 3) pelo usuário do sistema. No entanto, dada a descrição de alto nível da reverberação retardada, a sua relação com os parâmetros da FDN e a capacidade de modificar o seu comportamento, uma ampla variedade de métodos é prevista para controlar várias modalidades do processador de reverberação retardada baseado na FDN, incluindo (mas sem ficar a eles limitados) o que segue:

[00145] 1. O usuário final pode controlar manualmente os parâmetros da FDN, por exemplo, por meio de uma interface do usuário em um monitor (por exemplo, implementado por uma modalidade do subsistema de controle 209 da figura 3) ou de pré-ajustes de comutação ao usar controles físicos (por exemplo, implementado por uma modalidade do subsistema de controle 209 da figura 3). Desta maneira, o usuário final pode adaptar a simulação do ambiente de acordo com o gosto, o ambiente, ou o conteúdo;

[00146] 2. O autor do conteúdo de áudio a ser virtualizado pode prover os ajustes ou os parâmetros desejados que são conduzidos com o próprio conteúdo, por exemplo, pelos metadados providos com o sinal de áudio de entrada. Tais metadados podem ser analisados gramaticalmente e empregados (por exemplo, por uma modalidade do subsistema de controle 209 da figura 3) para controlar os parâmetros relevantes da FDN. Os metadados, portanto, podem ser indicativos de propriedades tais como o tempo de reverberação, o nível de reverberação, a razão direta a reverberação, e assim por diante, e essas propriedades podem ser variadas no tempo, sinalizadas pelos

metadados variados no tempo;

[00147] 3. O dispositivo de reprodução pode estar ciente de sua localização ou ambiente, por meio de um ou mais sensores. Por exemplo, um dispositivo móvel pode usar redes de GSM, sistema de posicionamento global (GPS), pontos de acesso WiFi conhecidos, ou qualquer outro serviço de localização para determinar onde o dispositivo se encontra. Subsequentemente, os dados indicativos da localização e/ou do ambiente podem ser empregados (por exemplo, por uma modalidade do subsistema de controle 209 da figura 3) para controlar os parâmetros relevantes da FDN. Desse modo, os parâmetros da FDN podem ser modificados em resposta à localização do dispositivo, por exemplo, para imitar o ambiente físico;

[00148] 4. Em relação à posição do dispositivo de reprodução, de um serviço de nuvem ou um meio social pode ser usado para derivar os ajustes mais comuns que os consumidores estão usando em um determinado ambiente. Além disso, os usuários podem fazer o upload de seus ajustes atuais para uma nuvem ou um serviço de meio social, em associação com a localização (conhecida) para tornar disponível para outros usuários, ou eles mesmos;

[00149] 5. O dispositivo de reprodução pode conter outros sensores tais como uma câmera, um sensor de luz, um microfone, um acelerômetro, um giroscópio, para determinar a atividade do usuário e o ambiente no qual o usuário se encontra, para otimizar os parâmetros da FDN para essa atividade e/ou ambiente particulares;

[00150] 6. Os parâmetros da FDN podem ser controlados pelo conteúdo de áudio. Os algoritmos de classificação de áudio, ou o conteúdo manualmente anotado podem indicar se os segmentos de áudio compreendem fala, música, efeitos sonoros, silêncio, e algo do gênero. Os parâmetros da FDN podem ser ajustados de acordo com tais etiquetas. Por exemplo, a razão direta a reverberação pode ser

reduzida para que o diálogo melhore a inteligibilidade do diálogo. Além disso, a análise de vídeo pode ser usada para determinar a localização de um segmento de vídeo atual, e os parâmetros da FDN podem ser ajustados por conseguinte para simular mais proximamente o ambiente mostrado no vídeo; e/ou

[00151] 7. O sistema de reprodução de estado sólido pode usar ajustes da FDN diferentes como um dispositivo móvel, por exemplo, os ajustes podem ser dependentes do dispositivo. Um sistema de estado sólido presente em uma sala de estar pode simular um cenário de sala de estar típico (razoavelmente reverberante) com fontes distantes, ao passo que um dispositivo móvel pode tornar o conteúdo mais próximo do ouvinte.

[00152] Algumas implementações do virtualizador da invenção incluem FDNs (por exemplo, uma implementação das FDN da figura 4) que são configuradas para aplicar retardos fracionários assim como retardo de amostra integral. Por exemplo, em uma de tais implementações um elemento de retardo fracionário é conectado em cada tanque de reverberação em série com uma linha de retardo que aplica o retardo integral igual a um número inteiro de períodos da amostra (por exemplo, cada elemento de retardo fracionário é posicionado depois ou então em série com uma das linhas de retardo). O retardo fracionário pode ser aproximado por uma mudança de fase (multiplicação complexa unitária) em cada faixa de frequência que corresponde a uma fração do período da amostra:  $f = \tau/T$ , onde  $f$  é a fração de retardo,  $\tau$  é o retardo desejado para a faixa, e  $T$  é o período da amostra para a faixa. É bem conhecido como se aplica retardo fracionário no contexto de aplicação de reverberação no domínio de QMF.

[00153] Em uma primeira classe de modalidades, a invenção é um método de virtualização de fones auriculares para a geração de um sinal

binaural em resposta a um conjunto de canais (por exemplo, cada um dos canais, ou cada um dos canais da faixa de frequência total) de um sinal de entrada de áudio de múltiplos canais, o qual inclui as etapas de: (a) aplicação de uma resposta de impulso de ambiente binaural (BRIR) a cada canal do conjunto (por exemplo, ao convolver cada canal do conjunto com uma BRIR que corresponde ao dito canal, nos subsistemas 100 e 200 da figura 3, ou nos subsistemas 12,..., 14 e 15 da figura 2), desse modo gerando sinais filtrados (por exemplo, as saídas dos subsistemas 100 e 200 da figura 3, ou as saídas dos subsistemas 12,..., 14 e 15 da figura 2), incluindo o uso de pelo menos uma rede de retardo de feedback (por exemplo, as FDNs 203, 204,..., 205 da figura 3) para aplicar uma reverberação retardada comum a uma mixagem descendente (por exemplo, uma mixagem descendente monofônica) dos canais do conjunto; e (b) combinação dos sinais filtrados (por exemplo, no subsistema 210 da figura 3, ou no subsistema que compreende os elementos 16 e 18 da figura 2) para gerar o sinal binaural. Tipicamente, um banco de FDNs é usado para aplicar a reverberação retardada comum à mixagem descendente (por exemplo, em que cada FDN aplica a reverberação retardada a uma faixa de frequência diferente). Tipicamente, a etapa (a) inclui uma etapa de aplicação a cada canal do conjunto de uma porção de "resposta direta e reflexão antecipada" de uma BRIR de um só canal para o canal (por exemplo, no subsistema 100 da figura 3 ou nos subsistemas 12,..., 14 da figura 2), e a reverberação retardada comum foi gerada para emular os macroatributos coletivos de porções de reverberação retardada de pelo menos algumas (por exemplo, todas) as BRIRs de um só canal.

[00154] Em modalidades típicas na primeira classe, cada uma das FDNs é implementada no domínio do filtro de espelho de quadratura complexo híbrido (HCQMF) ou no domínio do filtro de espelho de quadratura (QMF), e em algumas de tais modalidades os atributos

acústicos espaciais dependentes da frequência do sinal binaural são controlados (por exemplo, ao usar o subsistema de controle 209 da figura 3) mediante o controle da configuração de cada FDN empregada para aplicar a reverberação retardada. Tipicamente, uma mixagem descendente monofônica dos canais (por exemplo, a mixagem descendente gerada pelo subsistema 201 da figura 3) é usada como entrada para as FDNs para a renderização binaural eficiente do conteúdo de áudio do sinal de múltiplos canais. Tipicamente, o processo de mixagem descendente é controlado com base em uma distância da fonte para cada canal (isto é, a distância entre uma suposta fonte do conteúdo de áudio do canal e de uma suposta posição do usuário) e depende da manipulação das respostas diretas que correspondem às distâncias da fonte a fim de preservar a estrutura temporal e de nível de cada BRIR (isto é, cada BRIR determinada pelas porções de resposta direta e de reflexão antecipada de uma BRIR de um só canal para um canal, junto com a reverberação retardada comum para uma mixagem descendente incluindo o canal). Embora os canais a ser mixados de modo descendente possam ser alinhados no tempo e escalados de maneiras diferentes durante a mixagem descendente, o nível apropriado e a relação de sincronismo entre as porções de resposta direta, de reflexão antecipada, e de reverberação retardada comum da BRIR para cada canal devem ser mantidos. Em modalidades que usam um único banco de FDN para gerar a porção de reverberação retardada comum para todos os canais que são mixados de modo descendente (para gerar uma mixagem descendente), o ganho apropriado e o retardo precisam ser aplicados (a cada canal que é mixado de modo descendente) durante a geração da mixagem descendente.

[00155] As modalidades típicas nesta classe incluem uma etapa de ajuste (por exemplo, ao usar o subsistema de controle 209 da figura 3) dos coeficientes da FDN que correspondem aos atributos dependentes

da frequência (por exemplo, o tempo de deterioração de reverberação, a coerência interaural, a densidade modal e a razão direta a retardada). Isso permite uma melhor combinação de ambientes acústicos e umas saídas de som mais naturais.

[00156] Em uma segunda classe de modalidades, a invenção é um método para a geração de um sinal binaural em resposta a um sinal de entrada de áudio de múltiplos canais, mediante a aplicação de uma resposta de impulso de ambiente binaural (BRIR) a cada canal (por exemplo, ao convolver cada canal com uma BRIR correspondente) de um conjunto de canais de sinal de entrada (por exemplo, cada um dos canais de sinal de entrada ou cada canal da faixa de frequência total do sinal de entrada), incluindo: o processamento de cada canal do conjunto em um primeiro trajeto de processamento (por exemplo, implementado pelo subsistema 100 da figura 3 ou pelos subsistemas 12,..., 14 da figura 2) que é configurado para modelar, e aplicar a cada dito canal, uma porção de resposta direta e de reflexão antecipada (por exemplo, a EBRIR aplicada pelo subsistema 12, 14, ou 15 da figura 2) de uma BRIR de um só canal para o canal; e o processamento de uma mixagem descendente (por exemplo, uma mixagem descendente monofônica) dos canais do conjunto em um segundo trajeto de processamento (por exemplo, implementado pelo subsistema 200 da figura 3 ou pelo subsistema 15 da figura 2), em paralelo com o primeiro trajeto de processamento. O segundo trajeto de processamento é configurado para modelar, e aplicar à mixagem descendente, uma reverberação retardada comum (por exemplo, a LBRIR aplicada pelo subsistema 15 da figura 2). Tipicamente, a reverberação retardada comum emula macroatributos de porções de reverberação retardada de pelo menos algumas (por exemplo, todas) as BRIRs de um só canal. Tipicamente, o segundo trajeto de processamento inclui pelo menos uma FDN (por exemplo, uma FDN para cada uma de múltiplas faixas de frequência).

Tipicamente, uma monomixagem descendente é usada como entrada a todos os tanques de reverberação de cada FDN implementada pelo segundo trajeto de processamento. Tipicamente, são providos mecanismos (por exemplo, o subsistema de controle 209 da figura 3) para o controle sistemático de macroatributos de cada FDN a fim de melhor simular ambientes acústicos e produzir uma virtualização binaural de som mais natural. Uma vez que a maioria de tais macroatributos é dependente da frequência, cada FDN é tipicamente implementada no domínio do filtro de espelho de quadratura complexo híbrido (HCQMF), no domínio da frequência, domínio, ou em um outro domínio do banco de filtros, e uma FDN diferente é usada para cada faixa de frequência. Um benefício primário da implementação das FDNs em um domínio do banco de filtros consiste em permitir a aplicação de reverberação com propriedades dependentes da frequência de reverberação. Em várias modalidades, as FDNs são implementadas em qualquer um de uma ampla variedade de domínios do banco de filtros, ao usar qualquer um de uma variedade de banco de filtros, incluindo, mas sem ficar a eles limitados, filtros de espelho de quadratura (QMF), filtros de resposta de impulso finito (filtros FIR), filtros da resposta de impulso infinito (filtros IIR), ou filtros de cruzamento.

[00157] Algumas modalidades na primeira classe (e na segunda classe) implementam uma ou mais das seguintes características:

[00158] 1. Uma implementação de FDN do domínio do banco de filtros (por exemplo, o domínio do filtro de espelho de quadratura complexo híbrido) (por exemplo, a implementação de FDN da figura 4), ou a implementação de FDN do domínio de banco de filtros híbrido e a implementação de filtro de reverberação retardada do domínio do tempo (por exemplo, a estrutura descrita com referência à figura 8), que permite tipicamente o ajuste independente dos parâmetros e/ou dos ajustes da FDN para cada faixa de frequência (que permite o controle

simples e flexível de atributos acústicos dependentes da frequência), por exemplo, mediante a provisão da capacidade de variar os retardos do tanque de reverberação em faixas diferentes de modo a mudar a densidade modal como uma função da frequência;

[00159] 2. O processo de mixagem descendente específico, empregado para gerar (a partir do sinal de áudio de entrada de múltiplos canais) o sinal mixado de modo descendente (por exemplo, a mixagem descendente monofônica) processado no segundo trajeto de processamento, depende da distância da fonte de cada canal e da manipulação da resposta direta a fim de manter o nível apropriado e a relação de tempo entre as respostas direta e retardada;

[00160] 3. O filtro de passagem de alta e baixa (por exemplo, o APF 301 da figura 4) é aplicado no segundo trajeto de processamento (por exemplo, na entrada ou na saída de um banco de FDNs) para introduzir a diversidade de fase e a densidade de eco aumentada sem mudar o espectro e/ou o timbre de reverberação resultante;

[00161] 4. Retardos fracionários são implementados no trajeto de feedback de cada FDN em uma estrutura de múltiplas taxas avaliada complexa para superar os problemas relacionados aos retardos quantificados à grade do fator de amostra descendente;

[00162] 5. Nas FDNs, as saídas do tanque de reverberação são linearmente mixadas diretamente nos canais binaurais (por exemplo, pela matriz 312 da figura 4), ao usar os coeficientes de mixagem de saída que são ajustados com base na coerência interaural desejada em cada faixa de frequência. Opcionalmente, o mapeamento dos tanques de reverberação para os canais binaurais de saída é alternante através das faixas de frequência para obter um retardo equilibrado entre os canais binaurais. Também opcionalmente, fatores de normalização são aplicados às saídas do tanque de reverberação para equalizar os seus níveis enquanto são conservados o retardo fracionário e a potência total;

[00163] 6. O tempo de deterioração da reverberação dependente da frequência é controlado (por exemplo, ao usar o subsistema de controle 209 da figura 3) mediante o ajuste de combinações apropriadas dos ganhos e retardos do tanque de reverberação em cada faixa de frequência para simular ambientes reais;

[00164] 7. Um fator de escalonamento é aplicado (por exemplo, pelos elementos 306 e 309 da figura 4) por faixa de frequência (por exemplo, na entrada ou na saída do trajeto de processamento relevante), para:

[00165] controlar uma razão direta a retardada (DLR) dependente da frequência que combina com aquela de um ambiente real (um modelo simples pode ser usado para computar o fator de escalonamento requerido com base na DLR alvo e no tempo de deterioração da reverberação, por exemplo, T60);

[00166] prover a atenuação de baixa frequência para mitigar os artefatos de coleta excessivos; e/ou

[00167] aplicar a moldagem espectral de campo difuso às respostas da FDN;

[00168] 8. Modelos paramétricos simples são implementados (por exemplo, pelo subsistema de controle 209 da figura 3) para controlar atributos dependentes da frequência essenciais da reverberação retardada, tais como o tempo de deterioração da reverberação, a coerência interaural, e/ou a razão direta a retardada.

[00169] Em algumas modalidades (por exemplo, para as aplicações em que a latência do sistema é crítica e o retardo causado pelos bancos de filtros de análise e de síntese é proibitivo), as estruturas de FDN do domínio de banco de filtros de modalidades típicas do sistema da invenção (por exemplo, a FDN da figura 4 em cada faixa de frequência) são substituídas pelas estruturas de FDN implementadas no domínio do tempo (por exemplo, a FDN 220 da figura 10, que pode ser implementada tal como mostrado na figura 9). Em modalidades do

domínio do tempo do sistema da invenção, os subsistemas das modalidades do domínio de banco de filtros que aplicam um fator de ganho de entrada ( $G_{in}$ ), ganhos do tanque de reverberação ( $g_i$ ) e ganhos de normalização ( $1/|g_i|$ ) são substituídos por filtros do domínio do tempo (e/ou por elementos de ganho) a fim de permitir controles dependentes da frequência. A matriz de mixagem de saída de uma implementação do domínio de banco de filtros típica (por exemplo, a matriz de mixagem de saída 312 da figura 4) é substituída (em modalidades do domínio do tempo típicas) por um conjunto de saída de filtros do domínio do tempo (por exemplo, os elementos 500 a 503 da implementação da figura 11 do elemento 424 da figura 9). Ao contrário dos outros filtros de modalidades do domínio do tempo típicas, a resposta de fase desse conjunto de saída dos filtros é tipicamente crítica (porque a conservação da energia e a coerência interaural podem ser afetadas pela resposta de fase). Em algumas modalidades do domínio do tempo, o retardo do tanque de reverberação é variado (por exemplo, ligeiramente variado) de seus valores em uma implementação correspondente do domínio do banco de filtros (por exemplo, para evitar o compartilhamento do passo do banco de filtros como um fator comum).

[00170] A figura 10 é um diagrama de blocos de uma modalidade do sistema de virtualização de fones auriculares da invenção similar àquela da figura 3, exceto pelo fato que os elementos 202 a 207 do sistema da figura 3 são substituídos no sistema da figura 10 por uma única FDN 220 que é implementada no domínio do tempo (por exemplo, a FDN 220 da figura 10 pode ser implementada tal como a FDN da figura 9). Na figura 10, dois sinais do domínio do tempo (canais esquerdo e direito) são emitidos do subsistema de processamento de resposta direta e reflexão antecipada 100, e dois sinais do domínio do tempo (canais esquerdo e direito) são emitidos do subsistema de processamento de reverberação retardada 221. O elemento de adição 210 é acoplado às

saídas dos subsistemas 100 e 200. O elemento 210 é configurado para combinar (mixar) as saídas do canal esquerdo dos subsistemas 100 e 221 para gerar o canal esquerdo, L, do sinal de áudio binaural oriundo do virtualizador da figura 10, e para combinar (mixar) as saídas do canal direito dos subsistemas 100 e 221 para gerar o canal direito, R, do sinal áudio binaural oriundo do virtualizador da figura 10. O elemento 210 pode ser implementado para simplesmente somar a saída esquerda correspondente das amostras do canal dos subsistemas 100 e 221 para gerar o canal esquerdo do sinal de saída binaural, e simplesmente somar as amostras do canal direito correspondente dos subsistemas 100 e 221 para gerar o canal direito do sinal de saída binaural, supondo que os ajustes de nível e alinhamentos temporais apropriados são implementados nos subsistemas 100 e 221.

[00171] No sistema da figura 10, o sinal de entrada de áudio de múltiplos canais (que tem canais,  $X_i$ ) é dirigido para, e submetido a processamento em dois trajetos de processamento paralelos: um através do subsistema de processamento de resposta direta e reflexão antecipada 100; o outro através do subsistema de processamento de reverberação retardada 221. O sistema da figura 10 é configurado para aplicar uma  $BRIR_i$  a cada canal,  $X_i$ . Cada  $BRIR_i$  pode ser decomposta em duas porções: uma porção de resposta direta e de reflexão antecipada (aplicada pelo subsistema 100), e uma porção de reverberação retardada (aplicada pelo subsistema 221). Em operação, o subsistema processando de resposta direta e reflexão antecipada 100 gera desse modo as porções de resposta direta e de reflexão antecipada do sinal de áudio binaural que é emitido do virtualizador, e o subsistema de processamento de reverberação retardada ("gerador de reverberação retardada") 221 gera desse modo a porção de reverberação retardada do sinal de áudio binaural que é emitido do virtualizador. As saídas dos subsistemas 100 e 221 são mixadas (pelo

subsistema 210) para gerar o sinal de áudio binaural, o qual é aplicado tipicamente do subsistema 210 a um sistema de renderização (não mostrado) no qual é submetido a renderização binaural para a reprodução por fones auriculares.

[00172] O subsistema de mixagem descendente 201 (do subsistema de processamento de reverberação retardada 221) é configurado para mixar de modo descendente os canais de sinal de entrada de múltiplos canais em uma monomixagem descendente (que é o sinal do domínio do tempo), e a FDN 220 é configurada para aplicar a porção de reverberação retardada à monomixagem descendente.

[00173] Com referência à figura 9, será descrito em seguida um exemplo de uma FDN do domínio do tempo que pode ser empregada como a FDN 220 do virtualizador da figura 10. A FDN da figura 9 inclui o filtro de entrada 400, que é acoplado para receber uma monomixagem descendente (por exemplo, gerada pelo subsistema 201 do sistema da figura 10) de todos os canais de um sinal de entrada de áudio de múltiplos canais. A FDN da figura 9 também inclui o filtro de passagem de alta e baixa (APF) 401 (que corresponde ao APF 301 da figura 4) acoplado à saída do filtro 400, o elemento de ganho de entrada 401A acoplado à saída do filtro 401, 305, os elementos de adição 402, 403, 404 e 405 (que correspondem aos elementos de adição 302, 303, 304 e 305 da figura 4) acoplados à saída do elemento 401A, e quatro tanques de reverberação. Cada tanque de reverberação é acoplado à saída de um elemento diferente dos elementos 402, 403, 404 e 405, e compreende um dos filtros de reverberação 406 e 406A, 407 e 407A, 408 e 408A, e 409 e 409A, uma das linhas de retardo 410, 411, 412 e 413 (que correspondem às linhas de retardo 307 da figura 4) acopladas aos mesmos, e um dos elementos de ganho 417, 418, 419 e 420 acoplado à saída de uma das linhas de retardo.

[00174] A matriz unitária 415 (que corresponde à matriz unitária 308

da figura 4, e é tipicamente implementada para ser idêntica à matriz 308) é acoplada às saídas das linhas de retardo 410, 411, 412 e 413. A matriz 415 é configurada para aplicar uma saída de feedback a uma segunda entrada de cada um dos elementos 402, 403, 404 e 405.

[00175] Quando o retardo ( $n_1$ ) aplicado pela linha 410 é mais curto do que aquele ( $n_2$ ) aplicado pela linha 411, o retardo aplicado pela linha 411 é mais curto do que aquele ( $n_3$ ) aplicado pela linha 412, e o retardo aplicado pela linha 412 é mais curto do que aquele ( $n_4$ ) aplicado pela linha 413, as saídas dos elementos de ganho 417 e 419 (do primeiro e do terceiro tanques de reverberação) são aplicadas às entradas do elemento de adição 422, e as saídas dos elementos de ganho 418 e 420 (do segundo e do quarto tanques de reverberação) são aplicadas às entradas do elemento de adição 423. A saída do elemento 422 é aplicada a uma entrada de IACC e ao filtro de mixagem 424, e a saída do elemento 423 é aplicada à outra entrada do estágio de filtragem e mixagem 424 de IACC.

[00176] Os exemplos das implementações dos elementos do ganho 417 a 420 e dos elementos 422, 423 e 424 da figura 9 serão descritos com referência a uma implementação típica dos elementos 310 e 311 e da matriz de mixagem de saída 312 da figura 4. A matriz de mixagem de saída 312 da figura 4 (também identificada como matriz  $M_{out}$ ) é uma matriz de  $2 \times 2$  configurada para mixar os canais binaurais não mixados (as saídas dos elementos 310 e 311, respectivamente) da garimpagem inicial para gerar os canais de saída binaurais esquerdo e direito (os sinais da orelha esquerda, "L", e da orelha direita, "R", aplicados na saída da matriz 312) que têm a coerência interaural desejada. Essa garimpagem inicial é implementada pelos elementos 310 e 311, cada um dos quais combina duas saídas do tanque de reverberação para gerar um dos canais binaurais não mixados, em que a saída do tanque de reverberação tem o retardo mais curto aplicado a uma entrada do

elemento 310 e em que a saída do tanque de reverberação tem o segundo retardo mais curto aplicado a uma entrada do elemento 311. Os elementos 422 e 423 da modalidade da figura 9 executam o mesmo tipo de garimpagem inicial (nos sinais do domínio do tempo aplicados a suas entradas) ao passo que os elementos 310 e 311 (em cada faixa de frequência) da modalidade da figura 4 executam nas correntes dos componentes do domínio do banco de filtros (na faixa de frequência relevante) aplicadas a suas entradas.

[00177] Os canais binaurais não mixados (oriundos dos elementos 310 e 311 da figura 4, ou dos elementos 422 e 423 da figura 9), que estão perto de serem não correlacionados porque não consistem em nenhuma saída comum do tanque de reverberação, podem ser mixados (pela matriz 312 da figura 4 ou estágio 424 da figura 9) para implementar um padrão de garimpagem que provê uma coerência interaural desejada para os canais de saída binaurais esquerdo e direito. No entanto, devido ao fato que os retardos do tanque de reverberação são diferentes em cada FDN (isto é, a FDN da figura 9, ou a FDN implementada para cada faixa de frequência diferente na figura 4), um canal binaural não mixado (a saída de um dos elementos 310 e 311, ou 422 e 423) conduz constantemente a outro canal binaural não mixado (a saída do outro dos elementos 310 e 311, ou 422 e 423).

[00178] Desse modo, na modalidade da figura 4, se a combinação dos retardos do tanque de reverberação e do padrão de garimpagem for idêntica através de todas as faixas de frequência, deve resultar na polarização de som e imagem. Essa polarização pode ser mitigada se o padrão de garimpagem for alternado através das faixas de frequência de maneira tal que os canais de saída binaurais mixados conduzam e arrastem uns os outros em faixas de frequência alternadas. Por exemplo, se a coerência interaural desejada for Coh, onde  $ICohl \leq 1$ , a matriz de mixagem de saída 312 em faixas de frequência de números

ímpares pode ser implementada para multiplicar as duas entradas aplicadas à mesma por uma matriz que tem a seguinte forma:

$$M_{out} = \begin{bmatrix} \cos \beta & \sin \beta \\ \sin \beta & \cos \beta \end{bmatrix}, \quad \text{onde } \beta = \arcsin(Coh)/2,$$

e a matriz de mixagem de saída 312 em faixas de frequência de números pares pode ser implementada para multiplicar as duas entradas aplicadas à mesma por uma matriz que tem a seguinte forma:

$$M_{out,alt} = \begin{bmatrix} \sin \beta & \cos \beta \\ \cos \beta & \sin \beta \end{bmatrix}$$

onde  $\beta = \arcsin(Coh)/2$ .

[00179] Alternativamente, a polarização de som e imagem indicada acima nos canais de saída binaurais pode ser mitigada ao implementar para que a matriz 312 seja idêntica nas FDNs para todas as faixas de frequência, se a ordem do canal de suas entradas for comutada para alternar uma das faixas de frequência (por exemplo, a saída do elemento 310 pode ser aplicada à primeira entrada da matriz 312 e a saída do elemento 311 pode ser aplicada à segunda entrada da matriz 312 em faixas de frequência ímpares, e a saída do elemento 311 pode ser aplicada à primeira entrada da matriz 312 e a saída do elemento 310 pode ser aplicada à segunda entrada da matriz 312 nas faixas de frequência pares).

[00180] Na modalidade da figura 9 (e outras modalidades do domínio do tempo de uma FDN do sistema da invenção), não é trivial alternar a garimpagem com base na frequência para focar na polarização de som e imagem que deve então resultar quando o canal binaural não mixado oriundo do elemento 422 conduz constantemente (ou retarda) o canal binaural não mixado oriundo do elemento 423. Essa polarização de som e imagem é focada em uma modalidade típica do domínio do tempo de uma FDN do sistema da invenção de uma maneira diferente daquela que é tipicamente focada em uma modalidade do domínio do banco de

filtros de uma FDN do sistema da invenção. Especificamente, na modalidade da figura 9 (e algumas outras modalidades do domínio do tempo de uma FDN do sistema da invenção), os ganhos relativos dos canais binaurais não mixados (por exemplo, aqueles oriundos dos elementos 422 e 423 da figura 9) são determinados por elementos de ganho (por exemplo, os elementos 417, 418, 419 e 420 da figura 9) de modo a compensar a polarização de som e imagem que deve então resultar devido ao sincronismo desequilibrado observado. Com a implementação de um elemento de ganho (por exemplo, o elemento 417) para atenuar o sinal que chega primeiro (que foi garimpado em um lado, por exemplo, pelo elemento 422) e com a implementação de um elemento de ganho (por exemplo, o elemento 418) para impulsionar o sinal que chega em seguida (que foi garimpado no outro lado, por exemplo, pelo elemento 423), a imagem estéreo é recentrada. Desse modo, o tanque de reverberação incluindo o elemento de ganho 417 aplica um primeiro ganho à saída do elemento 417, e o tanque de reverberação incluindo o elemento de ganho 418 aplica um segundo ganho (diferente do primeiro ganho) à saída do elemento 418, de modo que o primeiro ganho e o segundo ganho atenuam o primeiro canal binaural não mixado (oriundo do elemento 422) em relação ao segundo canal binaural não mixado (oriundo do elemento 423).

[00181] Mais especificamente, em uma implementação típica da FDN da figura 9, as quatro linhas de retardo 410, 411, 412 e 413 têm um comprimento crescente, com os valores de retardo crescentes  $n_1$ ,  $n_2$ ,  $n_3$ , e  $n_4$ , respectivamente. Nesta implementação, o filtro 417 aplica outra vez  $g_1$ . Desse modo, a saída do filtro 417 é uma versão retardada da entrada para a linha de retardo 410 à qual um ganho de  $g_1$  foi aplicado. Similarmente, o filtro 418 aplica um ganho de  $g_2$ , o filtro 419 aplica um ganho de  $g_3$ , e o filtro 420 aplica um ganho de  $g_4$ . Desse modo, a saída do filtro 418 é uma versão retardada da entrada para a linha de

retardo 411 à qual um ganho de  $g_2$  foi aplicado, e a saída do filtro 419 é uma versão retardada da entrada para a linha de retardo 412 à qual um ganho de  $g_3$  foi aplicado, e a saída do filtro 420 é uma versão retardada da entrada para a linha de retardo 413 à qual um ganho de  $g_4$  foi aplicado.

[00182] Nesta implementação, a escolha dos seguintes valores de ganho pode resultar em uma polarização indesejável de som e imagem de saída (indicada pelos canais binaural oriundos do elemento 424) em um lado (isto é, ao canal esquerdo ou direito):  $g_1 = 0,5$ ,  $g_2 = 0,5$ ,  $g_3 = 0,5$  e  $g_4 = 0,5$ . De acordo com uma modalidade da invenção, os valores de ganho  $g_1$ ,  $g_2$ ,  $g_3$  e  $g_4$  (aplicados pelos elementos 417, 418, 419 e 420, respectivamente) são escolhidos tal como segue para centrar o som-imagem:  $g_1 = 0,38$ ,  $g_2 = 0,6$ ,  $g_3 = 0,5$  e  $g_4 = 0,5$ . Desse modo, a imagem estéreo de saída é recentrada de acordo com uma modalidade da invenção mediante a atenuação do sinal que chega primeiro (que foi garimpado em um lado, pelo elemento 422 no exemplo) em relação ao sinal que chega em penúltimo lugar (isto é, ao escolher  $g_1 < g_3$ ), e a impulsão do sinal que chega em seguida (que foi garimpado no outro lado, pelo elemento 423 no exemplo), em relação ao sinal que chega por último (isto é, ao escolher  $g_4 < g_2$ ).

[00183] As implementações típicas da FDN do domínio do tempo da figura 9 têm as seguintes diferenças e similaridades ao domínio do banco de filtros da FDN (domínio de CQMF) da figura 4:

[00184] a mesma matriz de feedback unitária, A (a matriz 308 da figura 4 e a matriz 415 da figura 9);

[00185] os retardos do tanque de reverberação similares,  $n_i$  (isto é, os retardos na implementação de CQMF da figura 4 podem ser  $n_1 = 17 \cdot 64T_s = 1088 \cdot T_s$ ,  $n_2 = 21 \cdot 64T_s = 1344 \cdot T_s$ ,  $n_3 = 26 \cdot 64T_s = 1664 \cdot T_s$ , e  $n_4 = 29 \cdot 64T_s = 1856 \cdot T_s$ , onde  $1/T_s$  é a taxa de amostra ( $1/T_s$  é tipicamente igual a 48 kHz), ao passo que os retardos na implementação

do domínio do tempo podem ser:  $n_1 = 1089 \cdot T_s$ ,  $n_2 = 1345 \cdot T_s$ ,  $n_3 = 1663 \cdot T_s$ , e  $n_4 = 185 \cdot T_s$ . Deve ser observado que nas implementações típicas de CQMF há uma restrição prática que cada retardo é algum múltiplo de número inteiro da duração de um bloco de 64 amostras (a taxa de amostra é tipicamente de 48 kHz), mas no domínio do tempo há mais flexibilidade quanto à escolha de cada retardo e, desse modo, mais flexibilidade quanto à escolha do retardo de cada tanque de reverberação);

[00186] implementações do filtro de passagem de alta e baixa similares (isto é, implementações similares do filtro 301 da figura 4 e do filtro 401 da figura 9). Por exemplo, o filtro de passagem de alta e baixa pode ser implementado mediante a colocação em cascata de vários filtros de passagem de alta e baixa (por exemplo, três). Por exemplo, cada filtro de passagem de alta e baixa na cascata pode ser da forma  $g - Z^{-n_i}/1 - g \cdot Z^{-n_i}$ , onde  $g = 0,6$ . O filtro de passagem de alta e baixa 301 da figura 4 pode ser implementado por três filtros de passagem de alta e baixa na cascata com retardos apropriados de blocos de amostra (por exemplo,  $n_1 = 64 \cdot T_s$ ,  $n_2 = 128 \cdot T_s$  e  $n_3 = 196 \cdot T_s$ ), ao passo que o filtro de passagem de alta e baixa 401 da figura 9 (o filtro de passa alta e baixa do domínio do tempo) pode ser implementado por três filtros de passagem de alta e baixa na cascata com retardos similares (por exemplo,  $n_1 = 61 \cdot T_s$ ,  $n_2 = 127 \cdot T_s$  e  $n_3 = 191 \cdot T_s$ ).

[00187] Em algumas implementações da FDN do domínio do tempo da figura 9, o filtro de entrada 400 é implementado de modo que faz com que a razão direta a retardada (DLR) da BRIR seja aplicada pelo sistema da figura 9 para combinar (pelo menos substancialmente) com uma DLR alvo, e de modo que a DLR da BRIR a ser aplicada por um virtualizador incluindo o sistema da figura 9 (por exemplo, o virtualizador da figura 10) possa ser mudada ao substituir o filtro 400 (ou ao controlar uma configuração do filtro 400). Por exemplo, em algumas modalidades,

o filtro 400 é implementado como uma cascata de filtros (por exemplo, um primeiro filtro 400A e um segundo filtro 400B, acoplados tal como mostrado na figura 9A) para implementar a DLR alvo e também implementar opcionalmente o controle da DLR desejado. Por exemplo, os filtros da cascata são filtros IIR (por exemplo, o filtro 400A é um filtro de passagem de alta frequência Butterworth de primeira ordem (um filtro IIR) configurado para combinar as características da baixa frequência alvo, e o filtro 400B é um filtro IIR de prateleira de baixa passagem de segunda ordem, configurado para combinar as características de alta frequência alvo). Para um outro exemplo, os filtros da cascata são filtros IIR e FIR (por exemplo, o filtro 400A é um filtro de passagem de alta frequência de Butterworth da segunda ordem (um filtro de IIR) configurado para combinar as características da frequência baixa do alvo, e o filtro 400B é um filtro FIR de 14ª ordem configurado para combinar as características de alta frequência alvo). Tipicamente, o sinal direto é fixo, e o filtro 400 modifica o sinal retardado para atingir a DLR alvo. O filtro de passagem de alta e baixa (APF) 401 é de preferência implementado para executar a mesma função que o APF 301 da figura 4, ou seja, para introduzir a diversidade de fase e a densidade de eco aumentada para gerar uma saída de FDN de som mais natural. O APF 401 controla tipicamente a resposta de fase, ao passo que o filtro de entrada 400 controla a resposta de amplitude.

[00188] Na figura 9, o filtro 406 e o elemento de ganho 406A implementam em conjunto um filtro de reverberação, o filtro 407 e o elemento de ganho 407A implementam em conjunto um outro filtro de reverberação, o filtro 408 e o elemento de ganho 408A implementam em conjunto um outro filtro de reverberação, e o filtro 409 e o elemento de ganho 409A implementam em conjunto um outro filtro de reverberação. Cada um dos filtros 406, 407, 408 e 409 da figura 9 é de preferência implementado como um filtro com um valor de ganho máximo perto de

um (ganho unitário), e cada um dos elementos de ganho 406A, 407A, 408A e 409A é configurado para aplicar um ganho de deterioração à saída de um filtro correspondente dos filtros 406, 407, 408 e 409 que combina com a deterioração desejada (depois do retardo do tanque de reverberação relevante,  $n_i$ ). Especificamente, o elemento de ganho 406A é configurado para aplicar um ganho de deterioração ( $\text{decaygain}_1$ ) à saída do filtro 406 para fazer com que a saída do elemento 406A tenha um ganho tal que a saída da linha de retardo 410 (depois do retardo do tanque de reverberação,  $n_1$ ) tenha um primeiro ganho deteriorado alvo, o elemento de ganho 407A é configurado para aplicar um ganho de deterioração ( $\text{decaygain}_2$ ) à saída do filtro 407 para fazer com a saída do elemento 407A tenha um ganho tal que a saída da linha de retardo 411 (depois do retardo do tanque de reverberação,  $n_2$ ) tenha um segundo ganho deteriorado alvo, o elemento de ganho 408A é configurado para aplicar um ganho de deterioração ( $\text{decaygain}_3$ ) à saída do filtro 408 para fazer com que a saída do elemento 408A tenha um ganho tal que a saída da linha de retardo 412 (depois do retardo do tanque de reverberação,  $n_3$ ) tenha um terceiro ganho deteriorado alvo, e o elemento de ganho 409A é configurado para aplicar um ganho de deterioração ( $\text{decaygain}_4$ ) à saída do filtro 409 para fazer com que a saída do elemento 409A tenha um ganho tal que a saída da linha de retardo 413 (depois do retardo do tanque de reverberação,  $n_4$ ) tenha um quarto ganho deteriorado alvo.

[00189] Cada um dos filtros 406, 407, 408 e 409, e cada um dos elementos 406A, 407A, 408A e 409A do sistema da figura 9 são implementados de preferência (em que cada um dos filtros 406, 407, 408 e 409 é de preferência implementado como um filtro IIR, por exemplo, um filtro de prateleira ou uma cascata de filtros de prateleira) para atingir um T60 alvo característico da BRIR a ser aplicada por um virtualizador incluindo o sistema da figura 9 (por exemplo, o virtualizador

da figura 10), onde "T60" denota o tempo de deterioração da reverberação ( $T_{60}$ ). Por exemplo, em algumas modalidades cada um dos filtros 406, 407, 408 e 409 é implementado como um filtro de prateleira (por exemplo, um filtro de prateleira que tem  $Q = 0,3$  e uma frequência de prateleira de 500 Hz, para atingir o T60 característico mostrado na figura 13, em que T60 tem unidades de segundos) ou como uma cascata de dois filtros de prateleira IIR (por exemplo, com frequências de prateleira de 100 Hz e 1.000 Hz, para atingir o T60 característico mostrado na figura 14, em que T60 tem unidades dos segundos). O formato de cada filtro de prateleira é determinado de modo a combinar com a curva de mudança desejada de baixa frequência a alta frequência. Quando o filtro 406 é implementado como um filtro da prateleira (ou uma cascata de filtros de prateleira), o filtro de reverberação que compreende o filtro 406 e o elemento de ganho 406A também é um filtro de prateleira (ou uma cascata de filtros de prateleira). Da mesma maneira, quando cada um dos filtros 407, 408 e 409 é implementado como um filtro de prateleira (ou uma cascata de filtros de prateleira), cada filtro de reverberação que compreende o filtro 407 (ou 408 ou 409) e o elemento de ganho correspondente (407A, 408A ou 409A) também é um filtro de prateleira (ou uma cascata de filtros de prateleira).

[00190] A figura 9B é um exemplo do filtro 406 implementado como uma cascata de um primeiro filtro de prateleira 406B e um segundo filtro de prateleira 406C, acoplados tal como mostrado na figura 9B. Cada um dos filtros 407, 408 e 409 pode ser implementado tal como é a implementação do filtro 406 da figura 9B.

[00191] Em algumas modalidades, os ganhos de deterioração ( $\text{decaygain}_i$ ) aplicados pelos elementos 406A, 407A, 408A e 409A são determinados tal como segue:

$$\text{decaygain}_i = 10^{((60 * (ni/Fs/T))/20)},$$

onde  $i$  é o índice do tanque de reverberação (isto é, o elemento 406A aplica  $\text{decaygain}_i$ , o elemento 407A aplica  $\text{decaygain}_2$ , e assim por diante),  $n_i$  é o retardo do  $i^{\circ}$  tanque de reverberação (por exemplo,  $n_1$  é o retardo aplicado pela linha de retardo 410),  $F_s$  é a taxa de amostragem,  $T$  é o tempo de deterioração da reverberação desejado ( $T_{60}$ ) a uma baixa frequência predeterminada.

[00192] a figura 11 é um diagrama de blocos de uma modalidade dos seguintes elementos da figura 9: elementos 422 e 423, e estágio de filtragem e mixagem 424 de IACC (coeficiente de correlação cruzada interaural). O elemento 422 é acoplado e configurado para somar as saídas dos filtros 417 e 419 (da figura 9) e para aplicar o sinal somado à entrada do filtro de prateleira de baixa passagem 500, e o elemento 422 é acoplado e configurado para somar as saídas dos filtros 418 e 420 (da figura 9) e para aplicar o sinal somado à entrada do filtro de alta passagem 501. As saídas dos filtros 500 e 501 são somadas (mixadas) no elemento 502 para gerar o sinal de saída binaural da orelha esquerda, e as saídas dos filtros 500 e 501 são mixadas no elemento 502 (a saída do filtro 500 é subtraída da saída do filtro 501) no elemento 502 para gerar o sinal de saída binaural da orelha direita. Os elementos 502 e 503 mixam (somam e subtraem) as saídas filtradas dos filtros 500 e 501 para gerar os sinais de saída binaurais que atingem (dentro da precisão aceitável) o IACC alvo característico. Na modalidade da figura 11, cada uma dentre o filtro de prateleira de baixa passagem 500 e o filtro elevado de alta passagem 501 é tipicamente implementado como um filtro IIR de primeira ordem. Em um exemplo no qual os filtros 500 e 501 têm tal implementação, a modalidade da figura 11 pode atingir o IACC característico exemplificador traçado como curva "I" na figura 12, que é uma boa combinação para o IACC característico alvo traçado como "I<sub>T</sub>" na figura 12.

[00193] A figura 11A é um gráfico da resposta de frequência ( $R_1$ ) de

uma implementação típica do filtro 500 da figura 11, da resposta de frequência (R2) de uma implementação típica do filtro 501 da figura 11, e da resposta dos filtros 500 e 501 conectados em paralela. É aparente a partir da figura 11A que a resposta combinada é desejavelmente suave através da faixa de 100 Hz a 10.000 Hz.

[00194] Desse modo, em uma classe de modalidades, a invenção é um sistema (por exemplo, aquele da figura 10) e o método para a geração de um sinal binaural (por exemplo, a saída do elemento 210 da figura 10) em resposta a um conjunto de canais de um sinal de entrada de áudio de múltiplos canais, que inclui a aplicação de uma resposta de impulso de ambiente binaural (BRIR) a cada canal do conjunto, desse modo gerando os sinais filtrados, incluindo o uso de uma única rede de retardo de feedback (FDN) para aplicar uma reverberação retardada comum a uma mixagem descendente dos canais do conjunto; e a combinação dos sinais filtrados para gerar o sinal binaural. A FDN é implementada no domínio do tempo. Em algumas tais modalidades, a FDN do domínio do tempo (por exemplo, a FDN 220 da figura 10, configurada tal como na figura 9) inclui:

[00195] um filtro de entrada (por exemplo, o filtro 400 da figura 9) que tem uma entrada acoplada para receber a mixagem descendente, em que o filtro de entrada é configurado para gerar uma primeira mixagem descendente filtrada em resposta à mixagem descendente;

[00196] um filtro de passagem de alta e baixa (por exemplo, o filtro de passagem de alta e baixa 401 da figura 9), acoplado e configurado a uma segunda mixagem descendente filtrada em resposta à primeira mixagem descendente filtrada;

[00197] um subsistema de aplicação de reverberação (por exemplo, todos os elementos da figura 9 com exceção dos elementos 400, 401 e 424), que tem uma primeira saída (por exemplo, a saída do elemento 422) e uma segunda saída (por exemplo, a saída do elemento 423), em

que o subsistema de aplicação de reverberação compreende um conjunto de tanques de reverberação, em que cada um dos tanques de reverberação tem um retardo diferente, e em que o subsistema de aplicação de reverberação é acoplado e configurado para gerar um primeiro canal binaural não mixado e um segundo canal binaural não mixado em resposta à segunda mixagem descendente filtrada, para aplicar o primeiro canal binaural não mixado na primeira saída, e para aplicar o segundo canal binaural não mixado na segunda saída; e

[00198] um estágio de filtragem e mixagem de coeficiente de correlação cruzada interaural (IACC) (por exemplo, o estágio 424 da figura 9, que pode ser implementado como elementos 500, 501, 502 e 503 de figura 11) acoplado ao subsistema de aplicação de reverberação e configurado para gerar um primeiro canal binaural mixado e um segundo canal binaural mixado em resposta ao primeiro canal binaural não mixado e a um segundo canal binaural não mixado.

[00199] O filtro de entrada pode ser implementado para gerar (de preferência como uma cascata de dois filtros configurados para gerar) a primeira mixagem descendente filtrada de maneira tal que cada BRIR tenha uma razão direta a retardada (DLR) que combina, pelo menos substancialmente, uma DLR alvo.

[00200] Cada tanque de reverberação pode ser configurado para gerar um sinal retardado, e pode incluir um filtro de reverberação (por exemplo, implementado como um filtro de prateleira ou uma cascata de filtros de prateleira) acoplado e configurado para aplicar um ganho a um sinal que se propaga em cada um dos ditos tanques de reverberação, para fazer com que o sinal retardado tenha um ganho que combine, pelo menos substancialmente, um ganho deteriorado alvo para o dito sinal retardado, em um esforço para obter um tempo de deterioração de reverberação alvo característico (por exemplo, um  $T_{60}$  característico) de cada BRIR.

[00201] Em algumas modalidades, o primeiro canal binaural não mixado conduz ao segundo canal binaural não mixado, os tanques de reverberação incluem um primeiro tanque de reverberação (por exemplo, o tanque de reverberação da figura 9 que inclui a linha de retardo 410) configurado para gerar um primeiro sinal retardado que tem um retardo mais curto e um segundo tanque de reverberação (por exemplo, o tanque de reverberação da figura 9 que inclui a linha de retardo 411) configurado para gerar um segundo sinal retardado que tem um segundo retardo mais curto, em que o primeiro tanque de reverberação é configurado para aplicar um primeiro ganho ao primeiro sinal retardado, o segundo tanque de reverberação é configurado para aplicar um segundo ganho ao segundo sinal retardado, em que o segundo ganho é diferente do primeiro ganho, ~~o segundo ganho é diferente do primeiro ganho,~~ e a aplicação do primeiro ganho e do segundo ganho resulta na atenuação do primeiro canal binaural não mixado em relação ao segundo canal binaural não mixado. Tipicamente, o primeiro canal binaural mixado e o segundo canal binaural mixado são indicativos de uma imagem estéreo recentrada. Em algumas modalidades, o estágio de filtragem e mixagem do IACC configurado para gerar o primeiro canal binaural mixado e o segundo canal binaural mixado de maneira tal que o dito primeiro canal binaural mixado e o dito segundo canal binaural mixado têm um IACC característico que combine pelo menos substancialmente um IACC alvo característica.

[00202] Os aspectos da invenção incluem os métodos e os sistemas (por exemplo, o sistema 20 da figura 2, ou o sistema da figura 3, ou da figura 10) que implementam (ou são configurados para implementar, ou suportam o desempenho de) a virtualização binaural de sinais de áudio (por exemplo, os sinais de áudio cujo conteúdo de áudio consiste em canais de alto-falante, e/ou sinais de áudio baseados em objetos).

[00203] Em algumas modalidades, o virtualizador da invenção é ou

inclui um processador de finalidade geral acoplado para receber ou gerar os dados de entrada indicativos de um sinal de entrada de áudio de múltiplos canais, e programado com um software (ou firmware) e/ou então configurado (por exemplo, em resposta aos dados de controle) para executar qualquer uma de uma variedade de operações nos dados de entrada, incluindo uma modalidade do método da invenção. Tal processador de finalidade geral deve ser acoplado tipicamente a um dispositivo de entrada (por exemplo, um mouse e/ou um teclado), a uma memória e a um dispositivo de exibição. Por exemplo, o sistema da figura 3 (ou o sistema 20 da figura 2, ou o sistema de virtualizador que compreende os elementos 12..., 14, 15, 16 e 18 do sistema 20) pode ser implementado em um processador de finalidade geral, em que as entradas são dados de áudio indicativos de N canais do sinal de entrada de áudio, e as saídas são dados de áudio indicativos de dois canais de um sinal de áudio binaural. Um conversor digital em analógico convencional (DAC) pode operar nos dados de saída para gerar versões analógicas dos canais de sinais binaurais para a reprodução pelos alto-falantes (por exemplo, um par de fones auriculares).

[00204] Embora as modalidades específicas da presente invenção e as aplicações da invenção tenham sido descritas no presente documento, será aparente aos elementos versados no estado da técnica que muitas variações nas modalidades e nas aplicações descritas no presente documento são possíveis sem desviar do âmbito da invenção descrita e reivindicada no presente documento. Deve ser compreendido que, embora determinadas formas da invenção tenham sido mostradas e descritas, a invenção não deve ser limitada às modalidades específicas descritas e mostradas ou aos métodos específicos descritos.

## REIVINDICAÇÕES

1. Método para gerar um sinal binaural em resposta a um conjunto de canais de um sinal de entrada de áudio de múltiplos canais, **caracterizado pelo fato de que** inclui as etapas de:

(a) aplicar uma resposta de impulso de ambiente binaural (BRIR) a cada canal do conjunto, desse modo gerando sinais filtrados, incluindo o uso de pelo menos uma rede de retardo de feedback para introduzir uma reverberação retardada comum em uma mixagem descendente dos canais do conjunto; e

(b) combinar os sinais filtrados para gerar o sinal binaural, em que, na etapa (a), a porção de reverberação retardada comum emula macroatributos coletivos das porções de reverberação retardada de BRIRs de canal único compartilhados através de pelo menos alguns canais do conjunto, em que os macroatributos coletivos compreendem uma ou mais dentre taxa de decaimento de reverberação, coerência interaural e distribuição espectral;

em que o método também inclui uma etapa de aplicação de valores de controle à rede de retardo de feedback para ajustar pelo menos um dentre ganho de entrada, ganhos do tanque de reverberação, retardos do tanque de reverberação ou parâmetros da matriz de saída para a rede de retardo de feedback, em que os valores de controle são aplicados de uma maneira tal que a porção de reverberação retardada comum emula os macroatributos coletivos das porções de reverberação retardada das BRIRs de canal único compartilhados através de pelo menos alguns canais do conjunto.

2. Método, de acordo com a reivindicação 1, **caracterizado pelo fato de que** a etapa (a) inclui uma etapa de geração de mixagem descendente de uma maneira que depende de uma distância da fonte para cada um dos canais que são mixados de modo descendente para gerar a mixagem descendente, e da manipulação de uma porção direta

da resposta da BRIR para cada um dos canais que são mixados de modo descendente para gerar a mixagem descendente, a fim de manter um nível apropriado e uma relação de sincronismo entre a porção de resposta da BRIR e a reverberação retardada comum.

3. Sistema configurado para gerar um sinal binaural em resposta a um sinal de entrada de áudio de múltiplos canais que tem canais, aplicando uma resposta de impulso de ambiente binaural a cada canal de um conjunto dos canais, em que o sistema é **caracterizado pelo fato de que** inclui:

um primeiro trajeto de processamento acoplado e configurado para aplicar, a cada canal do conjunto, pelo menos uma porção de resposta direta de uma resposta de impulso de ambiente binaural (BRIR) de canal único, para o canal; e

um segundo trajeto de processamento, acoplado em paralelo com o primeiro trajeto de processamento, e configurado para introduzir uma reverberação retardada comum em uma mixagem descendente dos canais do conjunto, em que a reverberação retardada comum emula macroatributos coletivos de porções de reverberação retardada de pelo menos de algumas das BRIRs de canal único através de pelo menos alguns canais do conjunto, em que os macroatributos coletivos compreendem uma ou mais dentre taxa de decaimento de reverberação, coerência interaural e distribuição spectral;

em que o segundo trajeto de processamento inclui pelo menos uma rede de retardo de feedback e o segundo trajeto de processamento é configurado para processar a mixagem descendente em pelo menos uma rede de retardo de feedback para introduzir a reverberação retardada comum na mixagem descendente,

em que o sistema também inclui:

um subsistema de controle acoplado e configurado para aplicar valores de controle à rede de retardo de feedback para ajustar

pelo menos um dentre ganho de entrada, ganhos do tanque de reverberação, retardos do tanque de reverberação ou parâmetros da matriz de saída para a dita rede de retardo de feedback, em que os valores de controle são aplicados de uma maneira tal que a porção de reverberação retardada comum emula macroatributos coletivos das porções de reverberação retardada das BRIRs de canal único compartilhadas através de pelo menos alguns canais do conjunto.

4. Sistema, de acordo com a reivindicação 3, **caracterizado pelo fato de que** o primeiro trajeto de processamento é configurado para gerar sinais filtrados em resposta a cada canal do conjunto, o segundo trajeto de processamento é configurado para gerar sinais filtrados adicionais em resposta à mixagem descendente, e em que o sistema também inclui:

um subsistema combinador de sinais, acoplado ao primeiro trajeto de processamento e ao segundo trajeto de processamento, e configurado para gerar o sinal binaural combinando os sinais filtrados e os sinais filtrados adicionais.

5. Sistema configurado para gerar um sinal binaural em resposta a um conjunto de canais de um sinal de entrada de áudio de múltiplos canais, em que o sistema é **caracterizado pelo fato de que** inclui:

um subsistema de filtragem acoplado e configurado para aplicar uma resposta de impulso de ambiente binaural (BRIR) a cada canal do conjunto, desse modo gerando sinais filtrados, incluindo a geração de uma mixagem descendente dos canais do conjunto e o processamento da mixagem descendente em pelo menos uma rede de retardo de feedback para introduzir uma reverberação retardada comum na mixagem descendente; e

um subsistema combinador de sinais, acoplado ao subsistema de filtragem, e configurado para gerar o sinal binaural

combinando os sinais filtrados,

em que a reverberação retardada comum emula macroatributos coletivos de porções de reverberação retardada BRIRs de canal único compartilhadas através de pelo menos alguns canais do conjunto, em que os macroatributos coletivos compreendem um ou mais dentre taxa de decaimento de reverberação, coerência interaural e distribuição spectral;

em que o sistema também inclui um subsistema de controle acoplado ao subsistema de filtragem e configurado para aplicar valores de controle à rede de retardo de feedback para ajustar pelo menos um dentre o ganho de entrada, ganhos do tanque de reverberação, retardos do tanque de reverberação ou parâmetros da matriz de saída para a rede de retardo de feedback, em que os valores de controle são aplicados de uma maneira tal que a porção de reverberação retardada comum emula macroatributos coletivos das porções de reverberação retardada das pelo menos algumas dentre as BRIRs de canal único compartilhadas através de pelo menos alguns canais do conjunto.

6. Sistema, de acordo com a reivindicação 5, **caracterizado pelo fato de que** o subsistema de filtragem é configurado para aplicar a cada canal do conjunto uma porção de resposta direta e de reflexão antecipada de uma BRIR de canal único para o canal.

7. Sistema, de acordo com a reivindicação 5, **caracterizado pelo fato de que** o subsistema de filtragem inclui um banco de redes de retardo de feedback configurado para aplicar a reverberação retardada comum à mixagem descendente, em que cada rede de retardo de feedback do banco introduz a reverberação retardada em uma faixa de frequência diferente da mixagem descendente.

8. Sistema, de acordo com a reivindicação 7, **caracterizado pelo fato de que** cada uma das redes de retardo de feedback é implementada no domínio do filtro de espelho de quadratura complexo.

9. Sistema, de acordo com a reivindicação 5, **caracterizado pelo fato de que** o sistema é um virtualizador de fones de ouvido.

10. Sistema, de acordo com a reivindicação 5, **caracterizado pelo fato de que** o sistema é um decodificador que inclui um subsistema virtualizador, e o subsistema virtualizador implementa o subsistema de filtragem e o subsistema de combinação de sinais.

11. Sistema, de acordo com a reivindicação 5, **caracterizado pelo fato de que** a mixagem descendente dos canais do conjunto é uma mixagem descendente monofônica dos canais do conjunto.

12. Sistema, de acordo com a reivindicação 5, **caracterizado pelo fato de que** o subsistema de filtragem inclui uma rede de retardo de feedback implementada no domínio do tempo, em que o subsistema de filtragem é configurado para processar a mixagem descendente no domínio do tempo na rede de retardo de feedback para introduzir a reverberação retardada comum na mixagem descendente.

13. Sistema, de acordo com a reivindicação 11, **caracterizado pelo fato de que** a rede de retardo de feedback inclui:

um filtro de entrada que tem uma entrada acoplada para receber a mixagem descendente, em que o filtro de entrada é configurado para gerar uma primeira mixagem descendente filtrado em resposta à mixagem descendente;

um filtro de passagem total, acoplado e configurado a uma segunda mixagem descendente filtrada em resposta à primeira mixagem descendente filtrada;

um subsistema de aplicação de reverberação, que tem uma primeira saída e uma segunda saída, em que o subsistema de aplicação de reverberação compreende um conjunto de tanques de reverberação, em que cada um dos tanques de reverberação tem um retardo diferente, e em que o subsistema de aplicação de reverberação é acoplado e configurado para gerar um primeiro canal binaural não mixado e um

segundo canal binaural não mixado em resposta à segunda mixagem descendente filtrada, para aplicar o primeiro canal binaural não mixado na primeira saída, e para aplicar o segundo canal binaural não mixado na segunda saída; e

um estágio de filtragem e mixagem de coeficiente de correlação cruzada interaural (IACC) acoplado ao subsistema de aplicação de reverberação e configurado para gerar um primeiro canal binaural mixado e um segundo canal binaural mixado em resposta ao primeiro canal binaural não mixado e a um segundo canal binaural não mixado.

14. Sistema, de acordo com a reivindicação 13, **caracterizado pelo fato de que** o filtro de entrada é implementado como uma cascata de dois filtros configurados para gerar a primeira mixagem descendente filtrada de maneira tal que cada BRIR tem uma razão direta a retardada (DLR) que combina, pelo menos substancialmente, com uma DLR alvo.

15. Sistema, de acordo com a reivindicação 13, **caracterizado pelo fato de que** cada um dos tanques de reverberação é configurado para gerar um sinal retardado, e inclui um filtro de reverberação acoplado e configurado para aplicar um ganho a um sinal que se propaga em cada um dos tanques de reverberação, para fazer com que o sinal retardado tenha um ganho que combine, pelo menos substancialmente, com um ganho deteriorado alvo para o sinal retardado, em um esforço para obter um tempo de deterioração de reverberação alvo característico de cada BRIR.

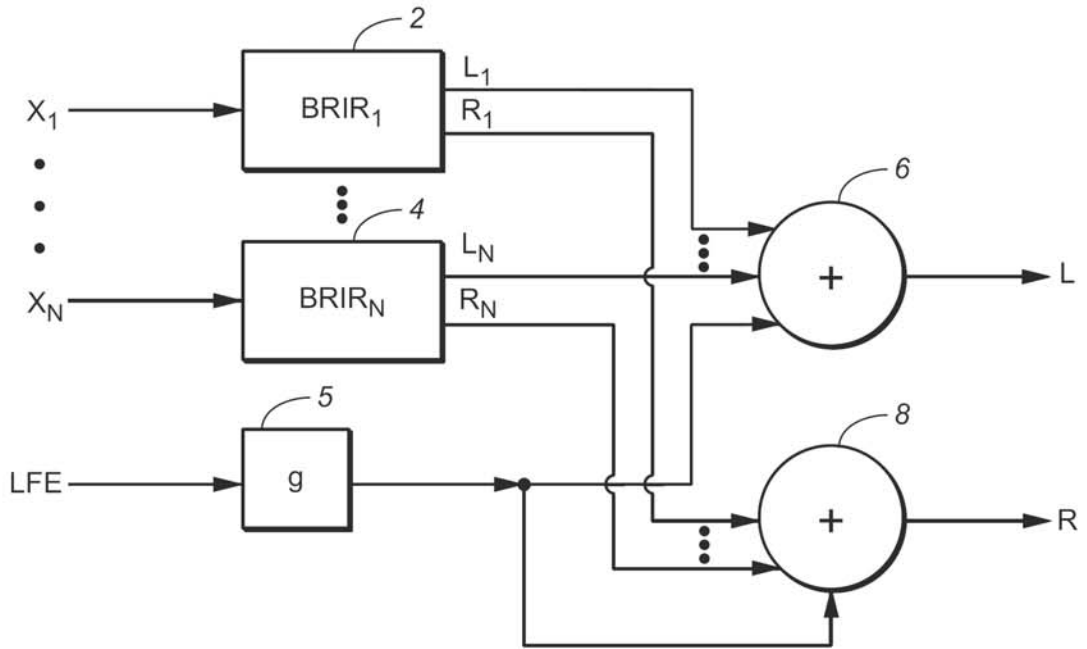
16. Sistema, de acordo com a reivindicação 15, **caracterizado pelo fato de que** cada filtro de reverberação é um filtro de prateleira ou uma cascata de filtros de prateleira.

17. Sistema, de acordo com a reivindicação 13, **caracterizado pelo fato de que** o primeiro canal binaural não mixado

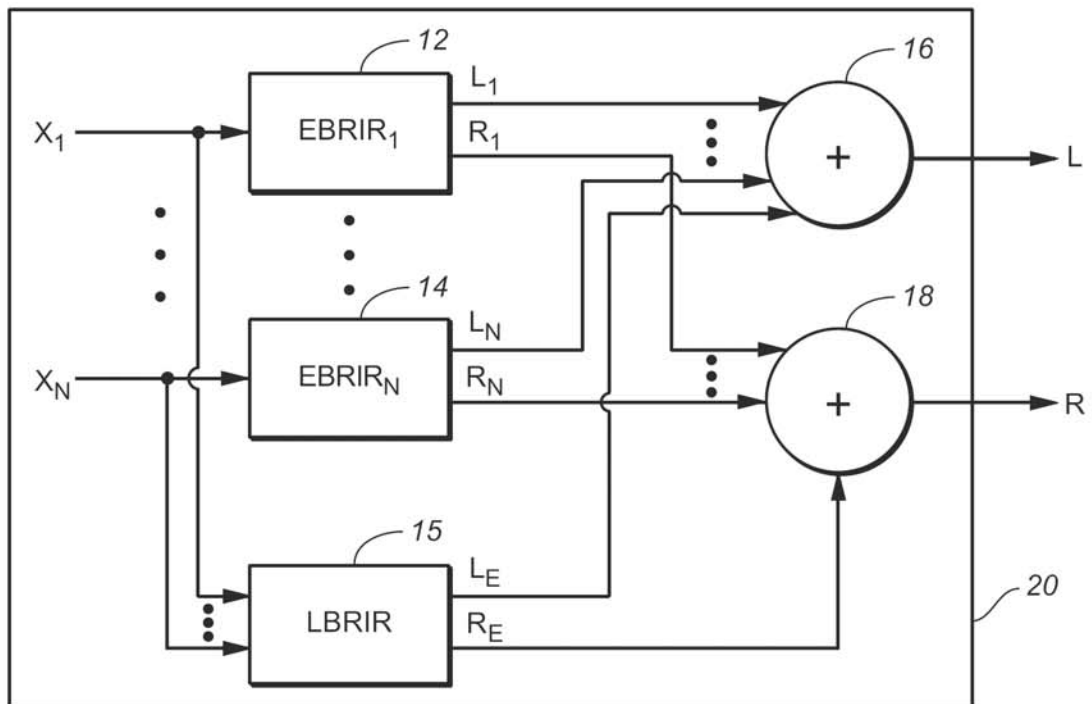
conduz o segundo canal binaural não mixado, em que os tanques de reverberação incluem um primeiro tanque de reverberação configurado para gerar um primeiro sinal retardado que tem um retardo mais curto e um segundo tanque de reverberação configurado para gerar um segundo sinal retardado que tem um segundo retardo mais curto, em que o primeiro tanque de reverberação é configurado para aplicar um primeiro ganho ao primeiro sinal retardado, o segundo tanque de reverberação é configurado para aplicar um segundo ganho ao segundo sinal retardado, o segundo ganho é diferente do primeiro ganho, e a aplicação do primeiro ganho e do segundo ganho resulta na atenuação do primeiro canal binaural não mixado em relação ao segundo canal binaural não mixado.

18. Sistema, de acordo com a reivindicação 13, **caracterizado pelo fato de que** o primeiro canal binaural mixado e o segundo canal binaural mixado são indicativos de uma imagem estéreo recentrada.

19. Sistema, de acordo com a reivindicação 13, **caracterizado pelo fato de que** o estágio de filtragem e mixagem do IACC é configurado para gerar o primeiro canal binaural mixado e o segundo canal binaural mixado de maneira tal que o primeiro canal binaural mixado e o segundo canal binaural mixado têm um IACC característico que combina pelo menos substancialmente com um IACC alvo característico.



**FIG. 1**  
(Técnica anterior)



**FIG. 2**

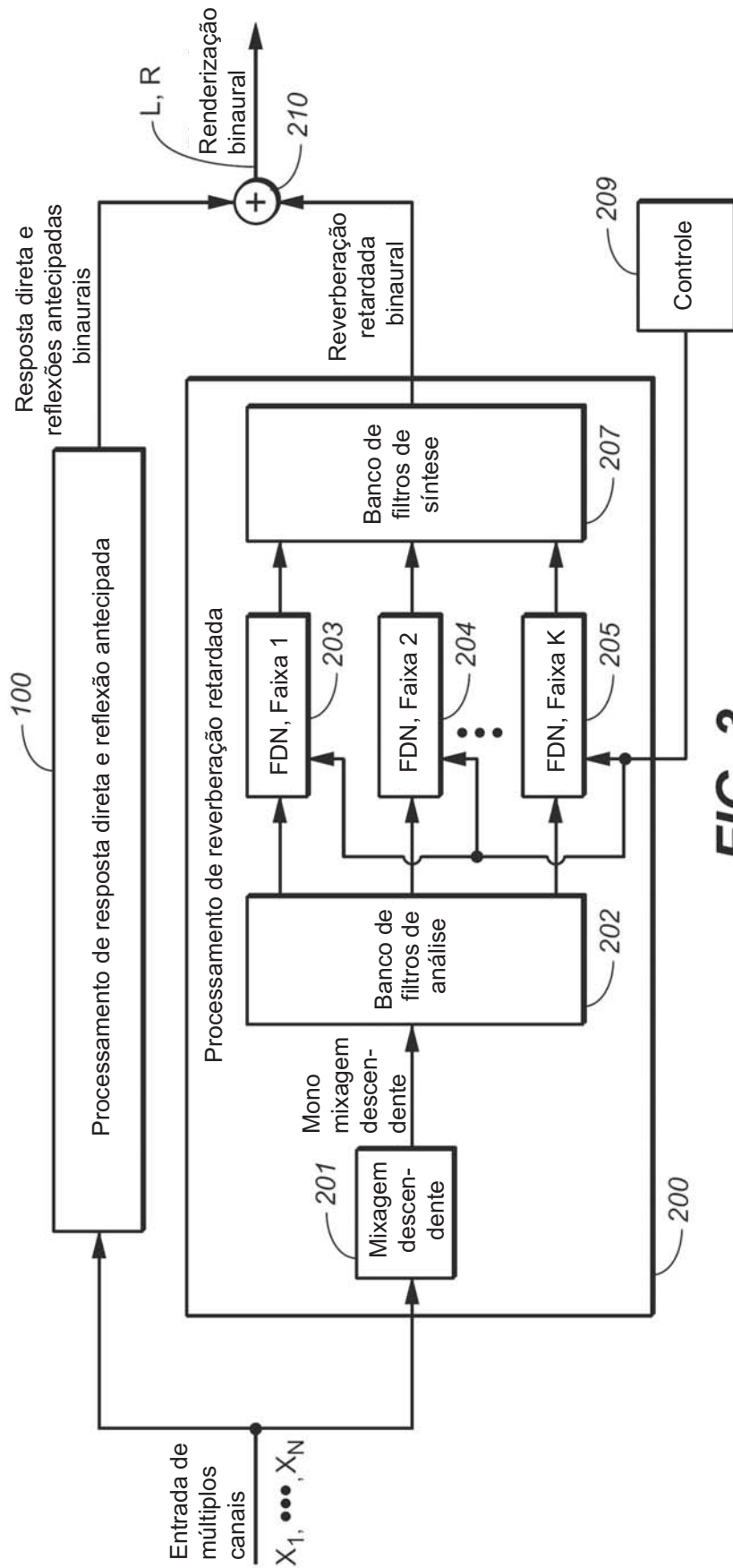
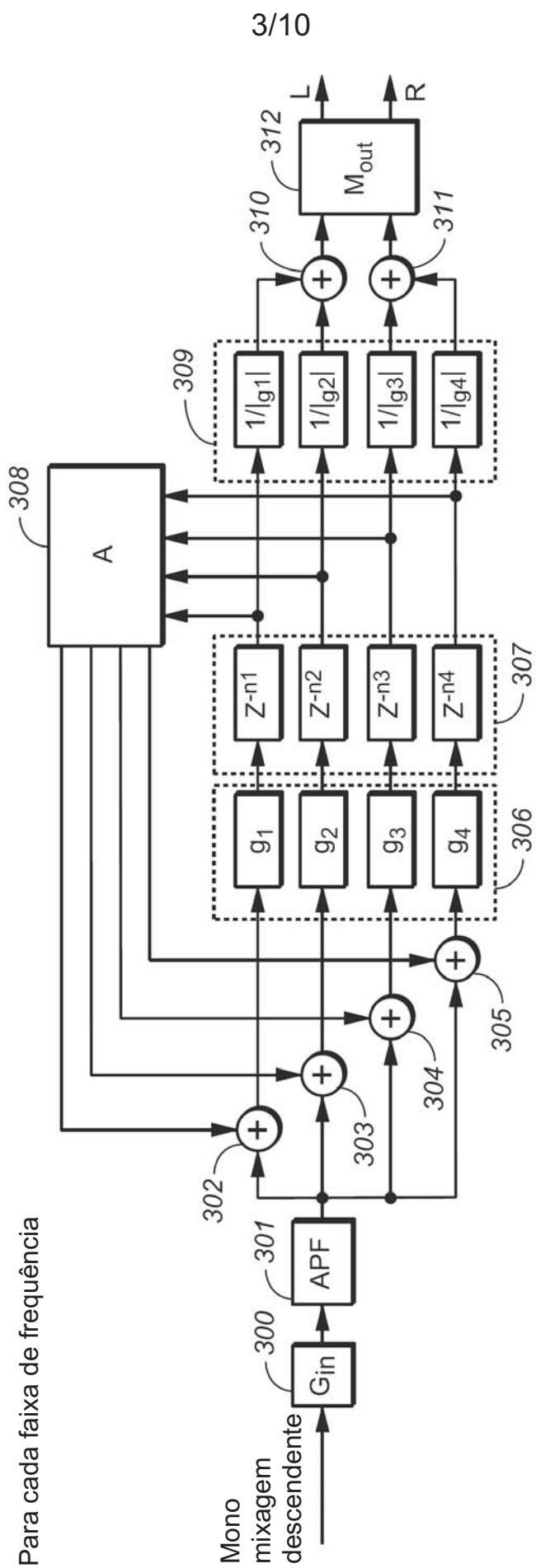
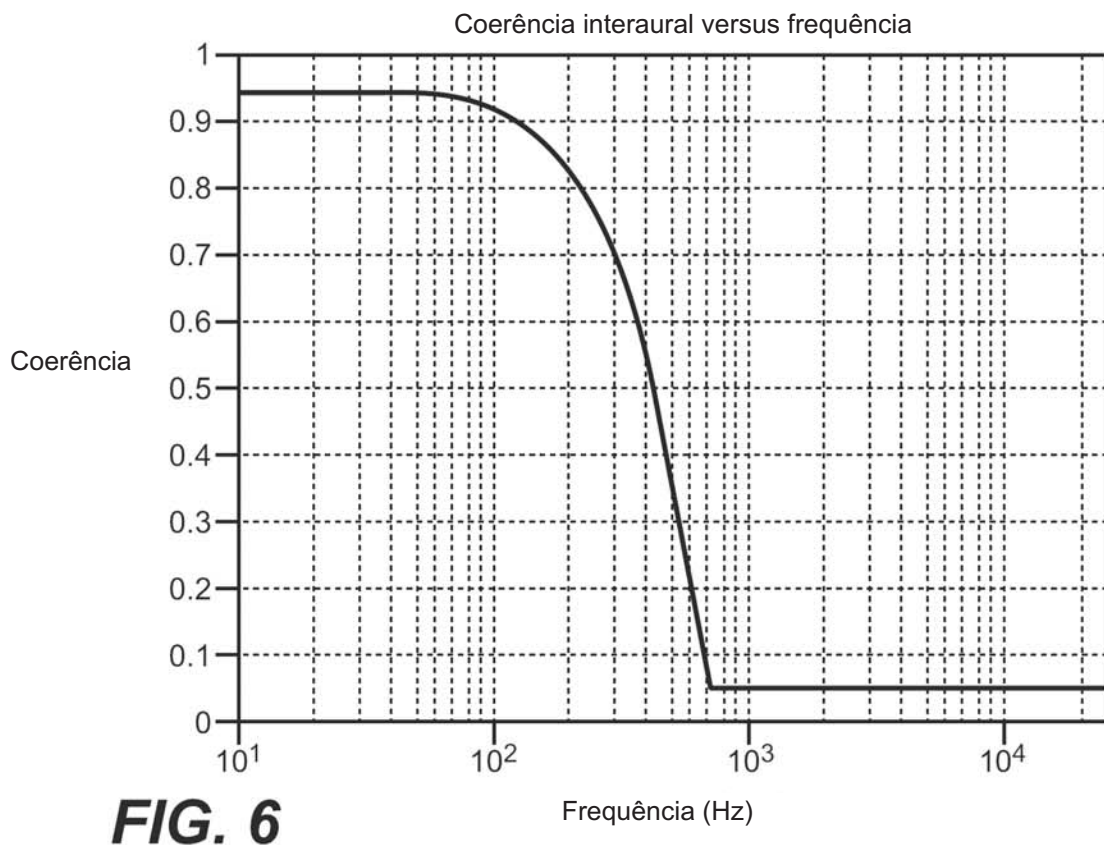
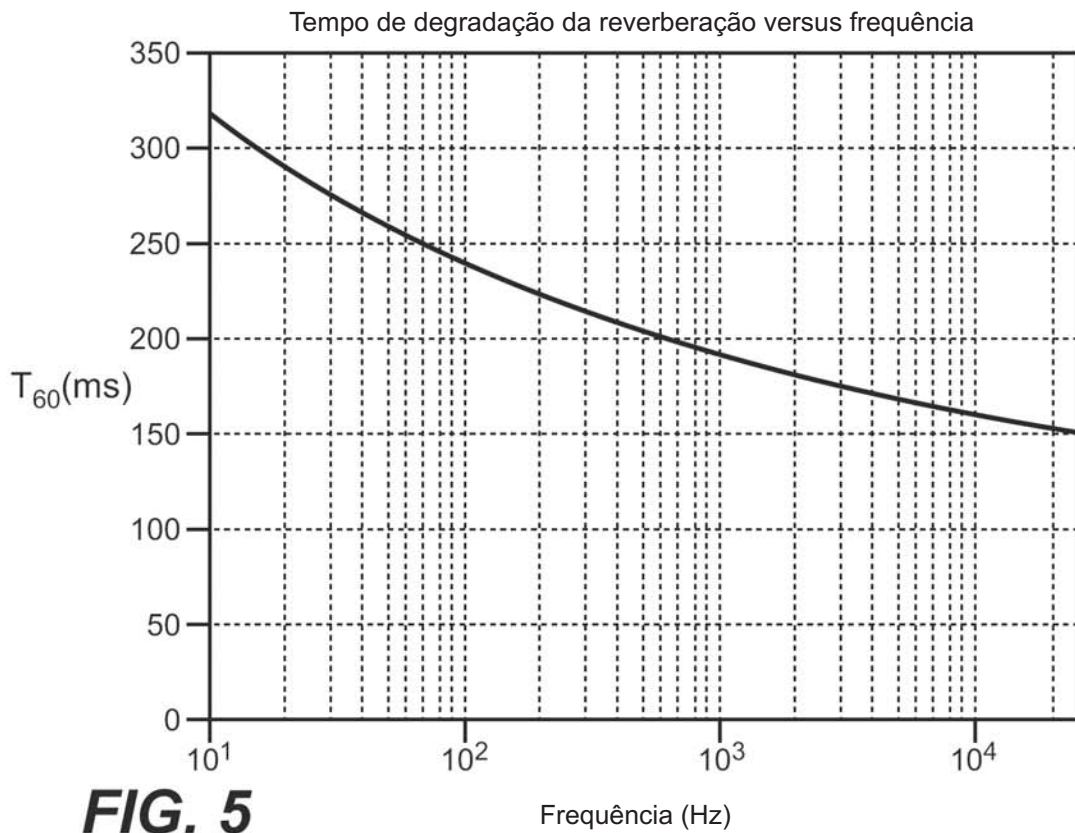
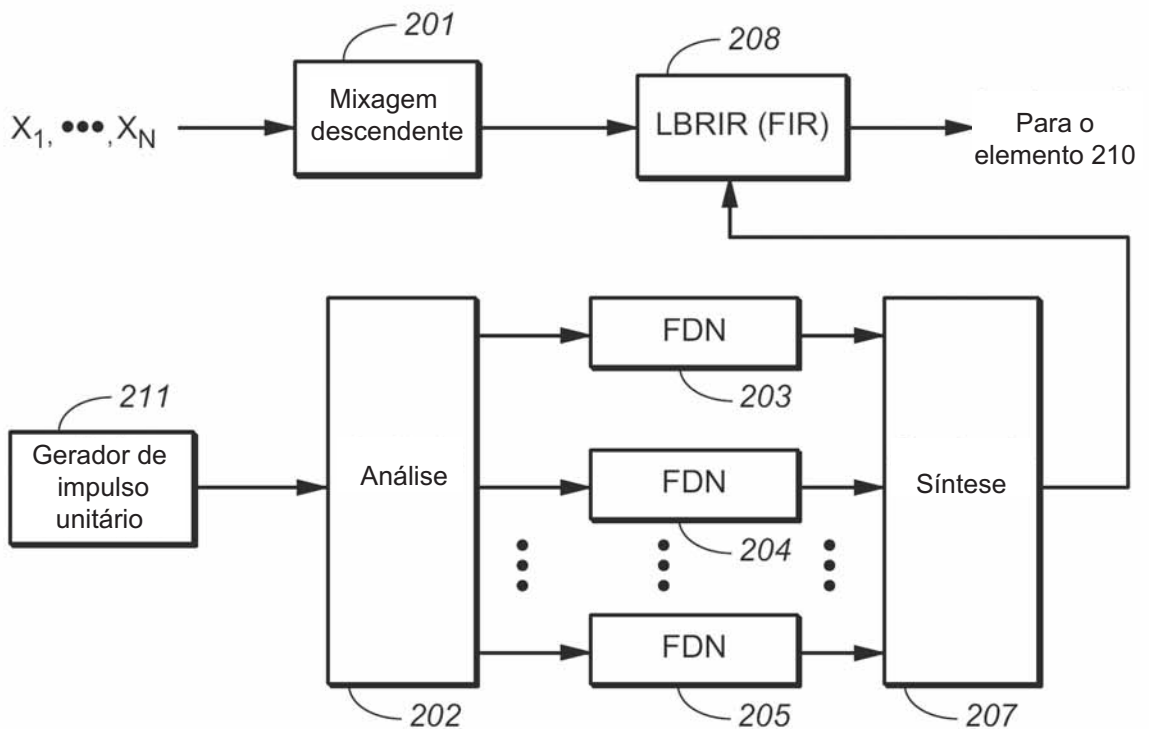
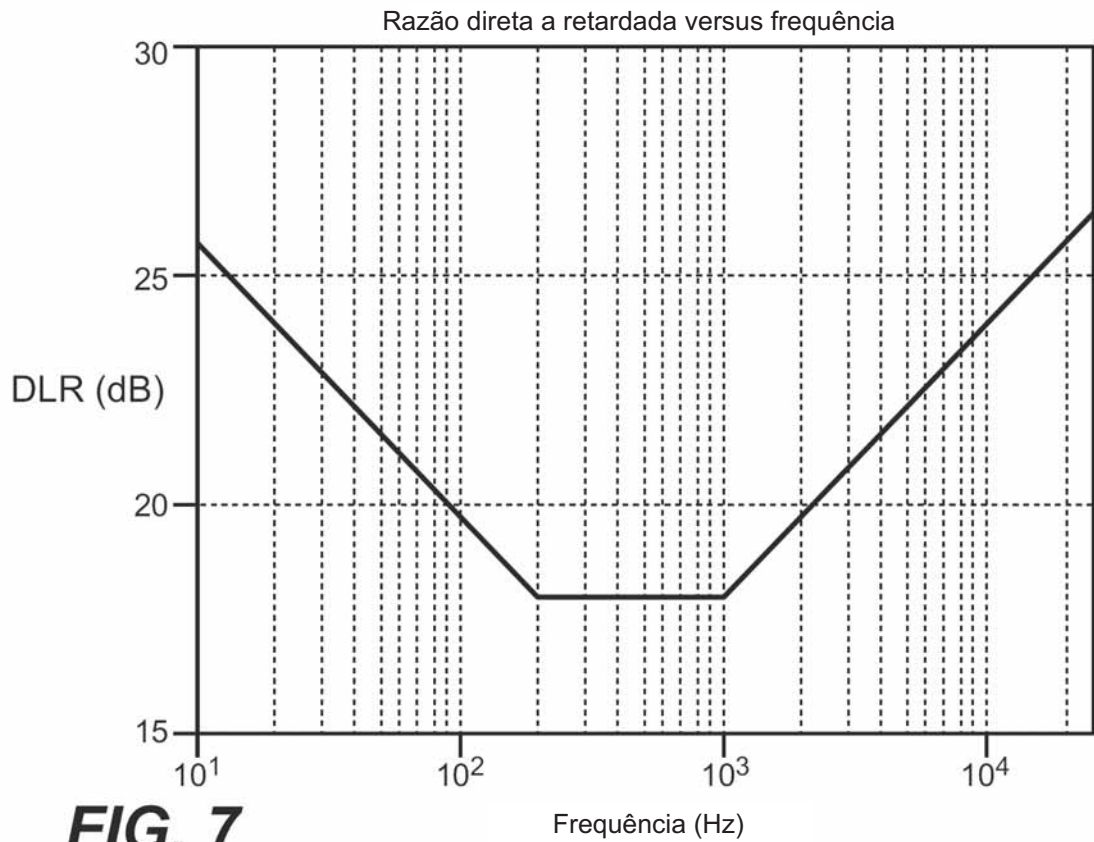


FIG. 3



**FIG. 4**





**FIG. 8**

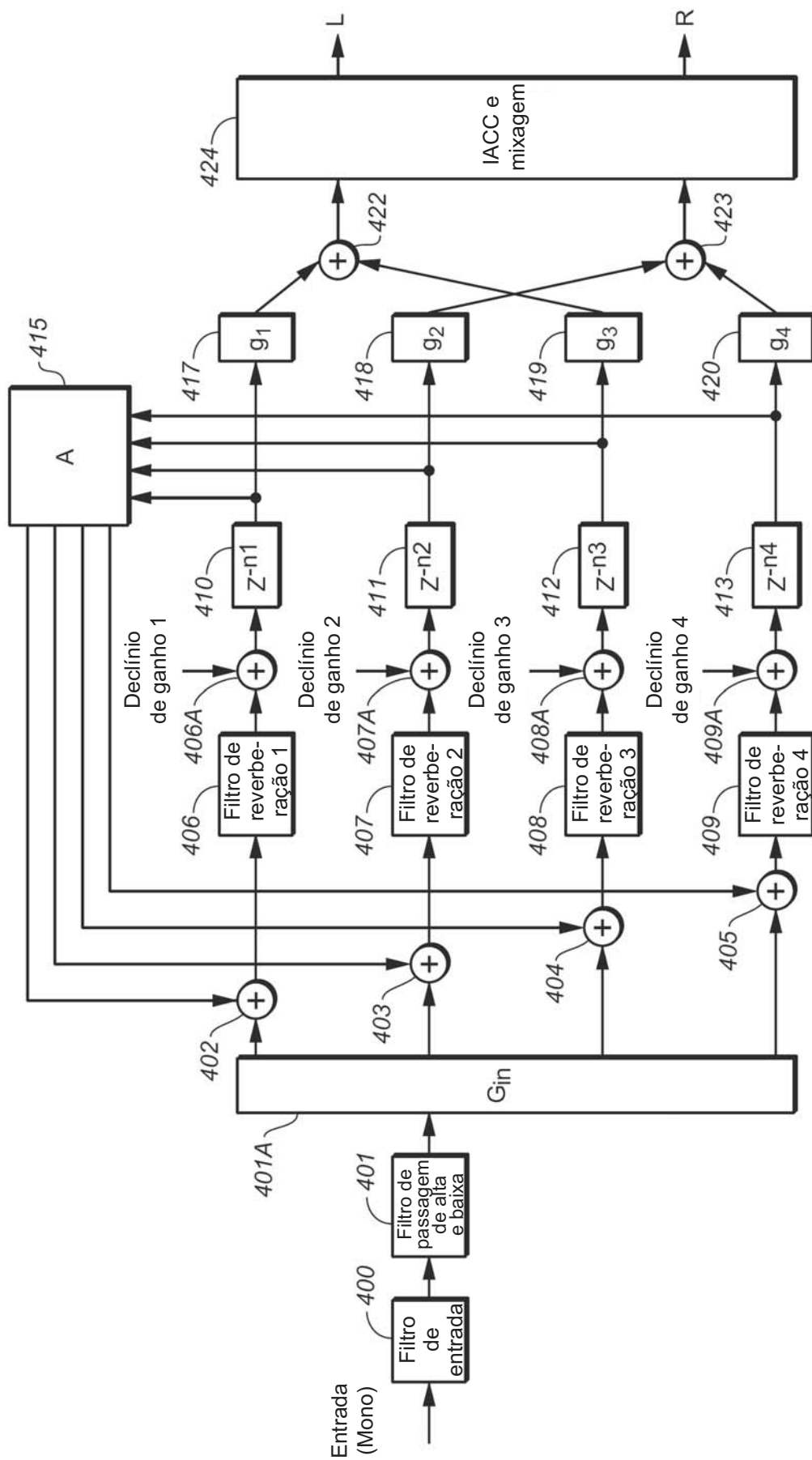
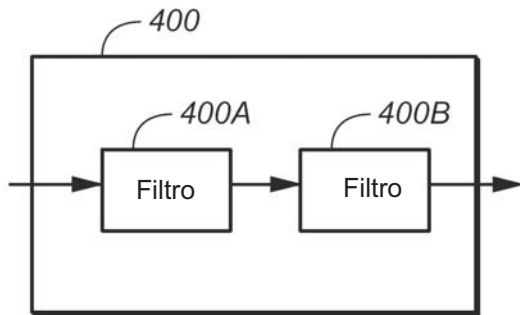
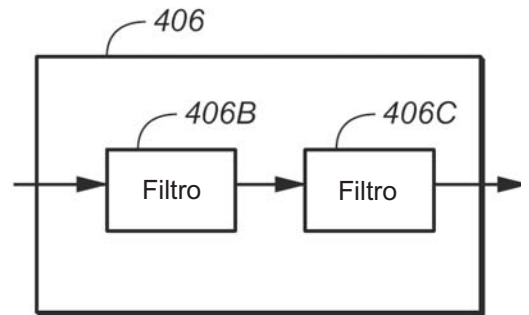
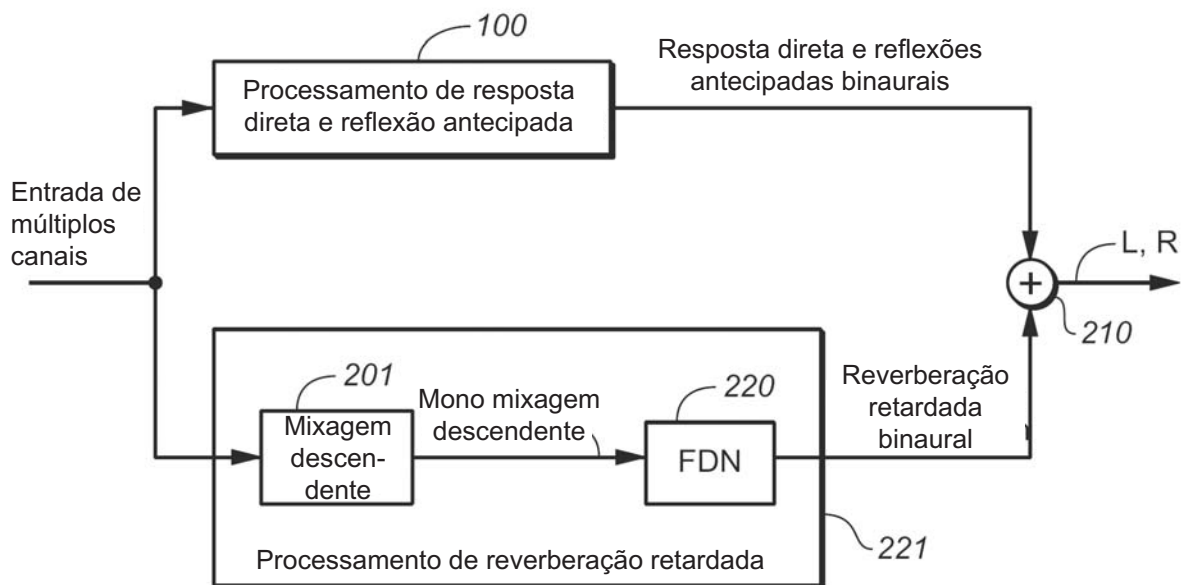
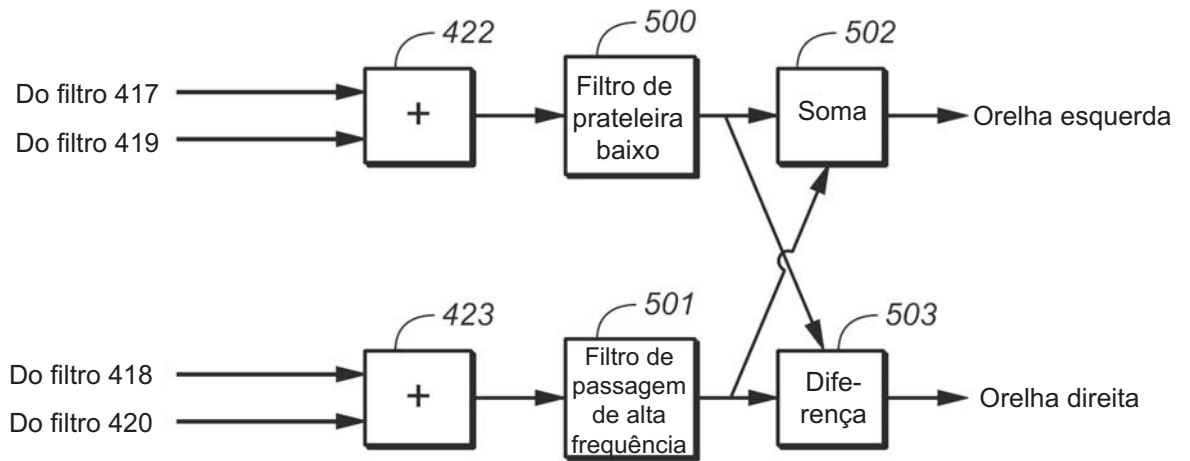
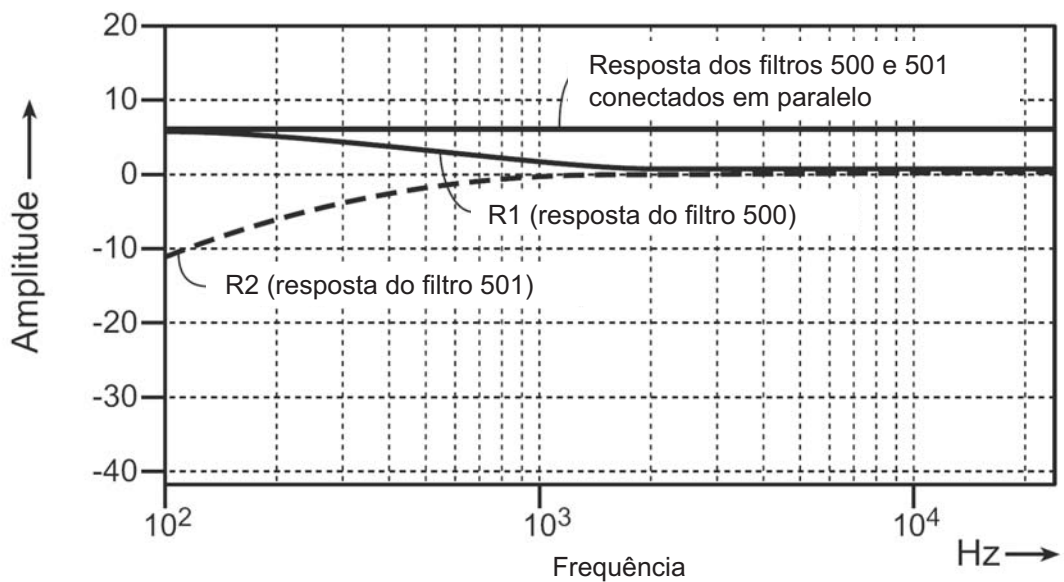
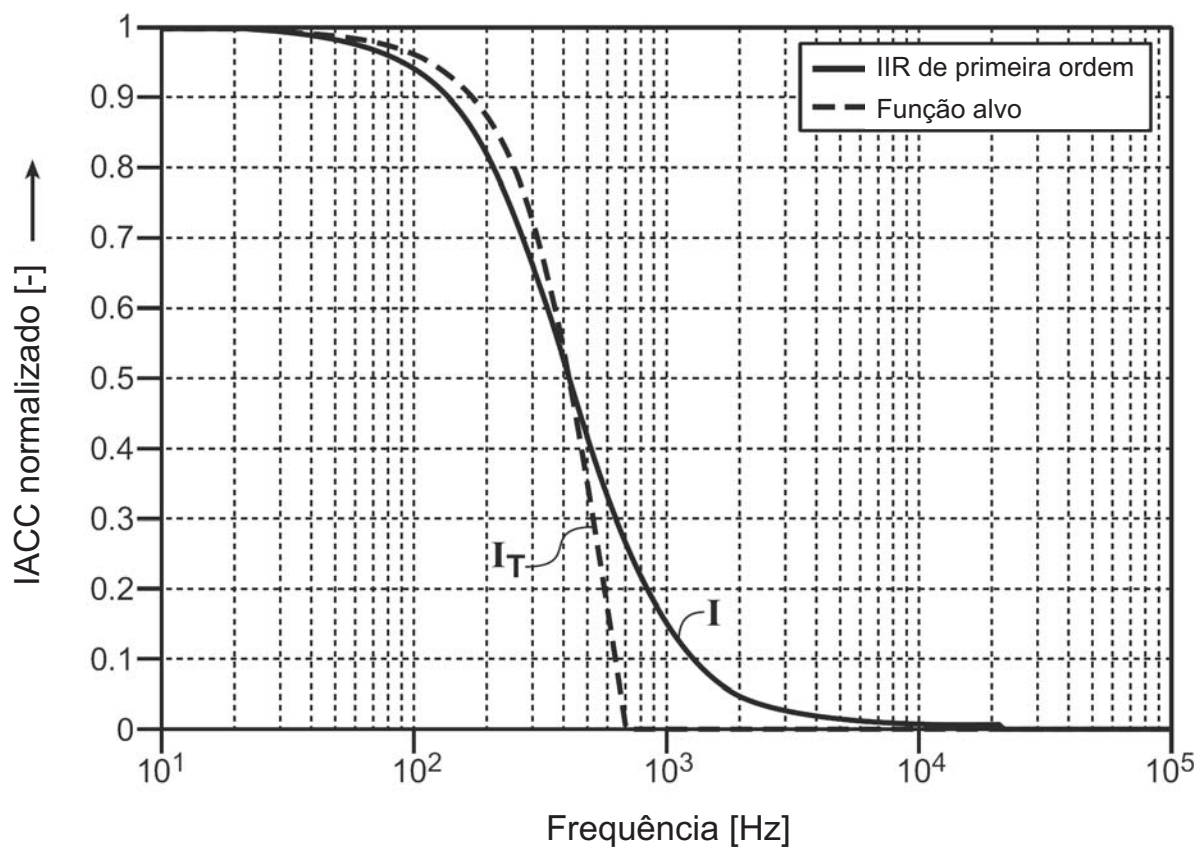
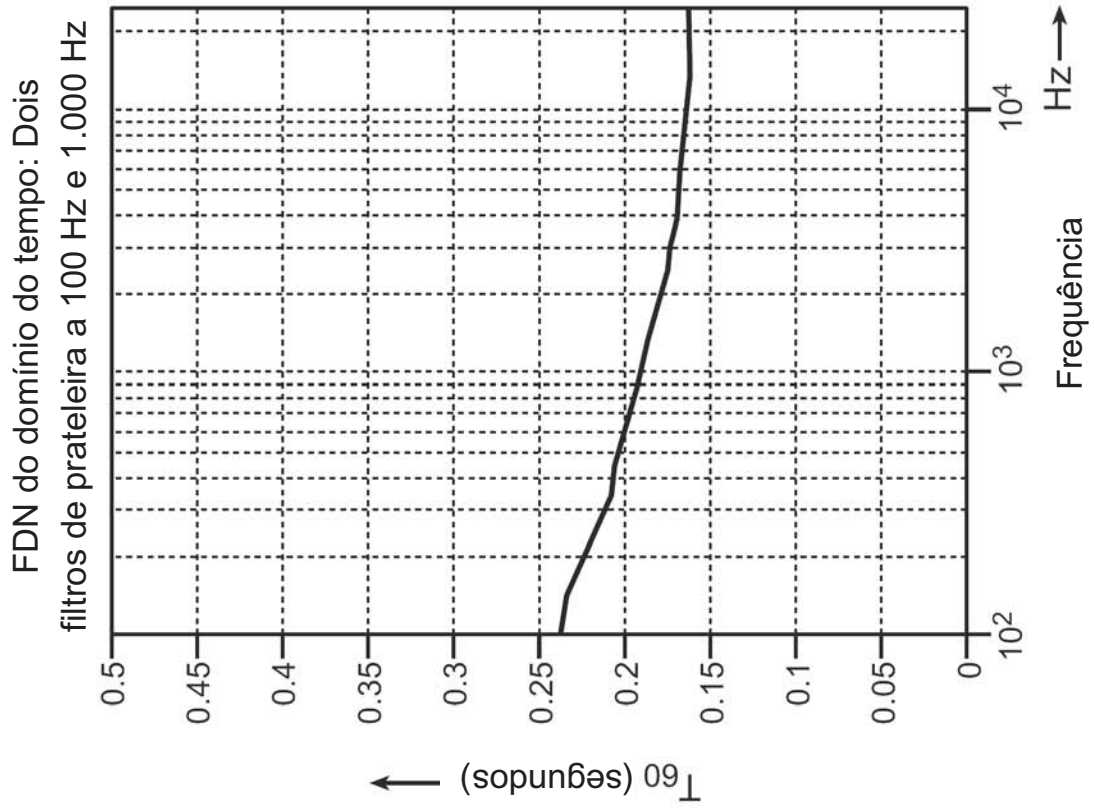


FIG. 9

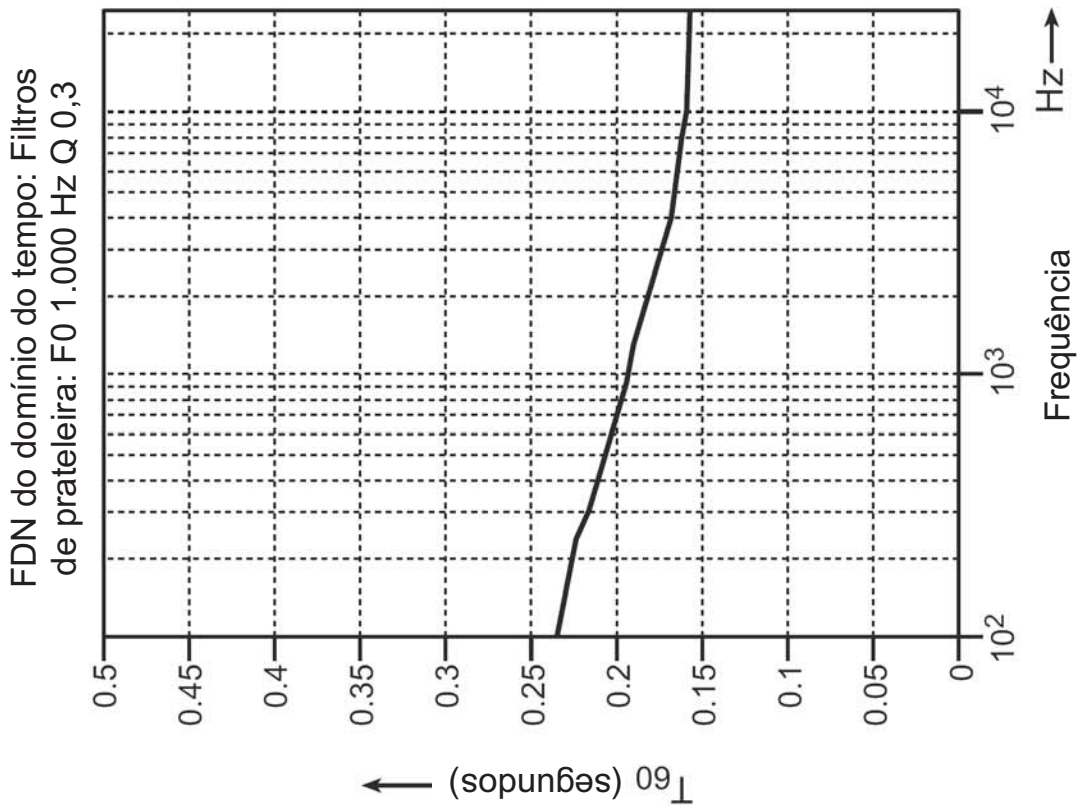
**FIG. 9A****FIG. 9B****FIG. 10**

**FIG. 11****FIG. 11A**

**FIG. 12**



**FIG. 14**



**FIG. 13**