



República Federativa do Brasil
Ministério do Desenvolvimento, Indústria
e do Comércio Exterior
Instituto Nacional da Propriedade Industrial

(21) **PI0706306-7 A2**

(22) Data de Depósito: 04/01/2007
(43) Data da Publicação: 22/03/2011
(RPI 2098)



(51) **Int.Cl.:**

H04S 1/00
H04R 5/00
G01L 19/00
G01L 19/02

(54) Título: **MÉTODO E APARELHO PARA A SINTETIZAÇÃO DE UM SINAL DE ÁUDIO BINAURAL; MÉTODO; MÉTODO PARA SINTETIZAÇÃO DE UM SINAL DE ÁUDIO ESTÉREO; DECODIFICADOR DE ÁUDIO PARAMÉTRICO; PRODUTO DE PROGRAMA DE COMPUTADOR, ARMAZENADO EM UMA MÍDIA LEGÍVEL POR COMPUTADOR E EXECUTÁVEL EM UM DISPOSITIVO DE PROCESSAMENTO DE DADOS, PARA PROCESSAR UM SINAL DE ÁUDIO PARAMETRICAMENTE CODIFICADO QUE COMPREENDE, AO MENOS, UM SINAL COMBINADO DE UMA PLURALIDADE DE CANAIS DE ÁUDIO E UM OU MAIS CONJUNTOS DE INFORMAÇÃO CORRESPONDENTES QUE DESCREVEM UMA IMAGEM SONORA DE CANAL MÚLTIPLO; MÉTODO PARA GERAR UM SINAL DE ÁUDIO PARAMETRICAMENTE CODIFICADO; CODIFICADOR DE ÁUDIO PARAMÉTRICO PARA GERAR UM SINAL ÁUDIO PARAMETRICAMENTE CODIFICADO; PRODUTO DE PROGRAMA DE COMPUTADOR, ARMAZENADO EM UMA MÍDIA LEGÍVEL POR COMPUTADOR E EXECUTÁVEL EM UM DISPOSITIVO DE PROCESSAMENTO DE DADOS, PARA GERAR UM SINAL DE ÁUDIO PARAMETRICAMENTE CODIFICADO**

(30) Prioridade Unionista: 17/01/2006 US 11/334,041,
09/01/2006 FI PCT/FI06/050014, 09/01/2006 FI PCT/FI06/050014,
17/01/2006 US 11/334,041

(73) Titular(es): Nokia Corporation

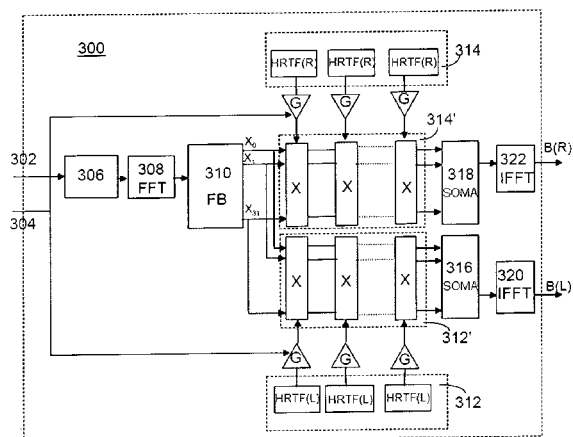
(72) Inventor(es): Julia Turku, Mauri Väänänen, Pasi Ojala

(74) Procurador(es): Araripe & Associados

(86) Pedido Internacional: PCT FI2007050004 de 04/01/2007

(87) Publicação Internacional: WO 2007/080224 de 19/07/2007

(57) Resumo: MÉTODO E APARELHO PARA A SINTETIZAÇÃO DE UM SINAL DE ÁUDIO BINAURAL; MÉTODO PARA SINTETIZAÇÃO DE UM SINAL DE ÁUDIO ESTÉREO; DECODIFICADOR DE ÁUDIO PARAMÉTRICO; PRODUTO DE PROGRAMA DE COMPUTADOR, ARMAZENADO EM UMA MÍDIA LEGÍVEL POR COMPUTADOR E EXECUTÁVEL EM UM DISPOSITIVO DE PROCESSAMENTO DE DADOS, PARA PROCESSAR UM SINAL DE ÁUDIO PARAMETRICAMENTE CODIFICADO QUE COMPREENDE, AO MENOS, UM SINAL COMBINADO DE UMA PLURALIDADE DE CANAIS DE ÁUDIO E UM OU MAIS CONJUNTOS DE INFORMAÇÃO CORRESPONDENTES QUE DESCREVEM UMA IMAGEM SONORA DE CANAL MÚLTIPLO; MÉTODO PARA GERAR UM SINAL DE ÁUDIO PARAMETRICAMENTE CODIFICADO; CODIFICADOR DE ÁUDIO PARAMÉTRICO PARA GERAR UM SINAL DE ÁUDIO PARAMETRICAMENTE CODIFICADO; PRODUTO DE PROGRAMA DE COMPUTADOR, ARMAZENADO EM UMA MÍDIA LEGÍVEL POR COMPUTADOR E EXECUTÁVEL EM UM DISPOSITIVO DE PROCESSAMENTO DE DADOS, PARA GERAR UM SINAL DE ÁUDIO PARAMETRICAMENTE CODIFICADO. Trata-se de um método para sintetizar um sinal de áudio binaural, sendo que o método compreende: inserir um sinal de áudio parametricamente codificado em, ao menos, um sinal combinado de uma pluralidade de canais de áudio e um ou mais conjuntos de informações secundárias que descrevem uma imagem sonora de canal múltiplo; e aplicar um conjunto pré-determinado de filtros de função de transferência relacionados à cabeça para, ao menos, um sinal combinado em proporção determinada pelo conjunto de informações secundárias correspondentes para sintetizar um sinal de áudio binaural. São descritos, ainda, um decodificador de áudio paramétrico, um codificador de áudio paramétrico, um produto de programa de computador e um aparelho para sintetização de um sinal de áudio binaural.



“MÉTODO E APARELHO PARA A SINTETIZAÇÃO DE UM SINAL DE ÁUDIO BINAURAL; MÉTODO PARA SINTETIZAÇÃO DE UM SINAL DE ÁUDIO ESTÉREO; DECODIFICADOR DE ÁUDIO PARAMÉTRICO; PRODUTO DE PROGRAMA DE COMPUTADOR, ARMAZENADO EM UMA MÍDIA LEGÍVEL POR COMPUTADOR E EXECUTÁVEL EM UM DISPOSITIVO DE PROCESSAMENTO DE DADOS, PARA PROCESSAR UM SINAL DE ÁUDIO PARAMETRICAMENTE CODIFICADO QUE COMPREENDE, AO MENOS, UM SINAL COMBINADO DE UMA PLURALIDADE DE CANAIS DE ÁUDIO E UM OU MAIS CONJUNTOS DE INFORMAÇÃO CORRESPONDENTES QUE DESCRIVEM UMA IMAGEM SONORA DE CANAL MÚLTIPLO; MÉTODO PARA GERAR UM SINAL DE ÁUDIO PARAMETRICAMENTE CODIFICADO; CODIFICADOR DE ÁUDIO PARAMÉTRICO PARA GERAR UM SINAL DE ÁUDIO PARAMETRICAMENTE CODIFICADO; PRODUTO DE PROGRAMA DE COMPUTADOR, ARMAZENADO EM UMA MÍDIA LEGÍVEL POR COMPUTADOR E EXECUTÁVEL EM UM DISPOSITIVO DE PROCESSAMENTO DE DADOS, PARA GERAR UM SINAL DE ÁUDIO PARAMETRICAMENTE CODIFICADO”

Pedidos Relacionados

Este pedido reivindica prioridade de um pedido internacional nos termos do PCT/FI2006/050014, depositado no dia 9 de janeiro de 2006 e de um pedido n° U.S. 11/334.041, depositado no dia 17 de janeiro de 2006.

Campo da Invenção

A presente invenção refere-se à codificação de áudio espacial e, mais particularmente, à decodificação de sinais de áudio binaurais.

Antecedentes da Invenção

Em codificação de áudio especial, um sinal de áudio de canal duplo ou múltiplo é processado de forma que os sinais de áudio a serem reproduzidos em diferentes canais de áudio sejam diferentes uns dos outros, fornecendo, assim, aos ouvintes, a impressão de um efeito espacial ao redor da fonte de áudio. O efeito espacial pode ser criado através da gravação do áudio diretamente em formatos adequados para reprodução

de canal múltiplo ou binaural ou o efeito espacial pode ser criado artificialmente em qualquer sinal de áudio de canal duplo ou múltiplo, o que é conhecido como espacialização.

5 Sabe-se, geralmente, que, para reprodução por fones de ouvido, uma espacialização artificial pode ser realizada por filtragem de HRTF (Função de Transferência Relativa à Cabeça), a qual produz sinais binaurais para o ouvido esquerdo e direito do ouvinte. Os sinais de fonte de som são filtrados com filtros derivados das HRTFs correspondentes a sua direção de origem. Uma HRTF é a função de transferência medida a partir de uma fonte de som em campo livre à audição humana ou uma cabeça
10 artificial dividida pela função de transferência para um microfone que substitui a cabeça e é situado no meio da cabeça. O efeito artificial de sala (por exemplo, reflexões iniciais e/ou reverberação tardia) pode ser adicionado aos sinais de espacialização para melhorar a externalização e a naturalidade da fonte.

Enquanto uma variedade de dispositivos de interação e de escuta de áudio
15 aumenta, compatibilidade se torna mais importante. Dentre os formatos de áudio espacial, aspira-se à compatibilidade através de técnicas de aumento de canais e redução de canais. Sabe-se, geralmente, que há algoritmos para conversão de um sinal de áudio de canal múltiplo para formato estéreo, como Dolby Digital® e Dolby Ambiente® e para uma conversão adicional de um sinal estéreo em sinal binaural. Entretanto, neste tipo de
20 processo, a imagem espacial do sinal de áudio de canal múltiplo original não pode ser completamente reproduzida. Uma melhor maneira de converter um sinal de áudio de canal múltiplo para escuta por fone de ouvido é substituir os alto-falantes originais empregando-se a filtragem de HRTF e executar os sinais de canal do alto-falante através daqueles (por exemplo, Dolby Headphone®). Porém, o dito processo tem a desvantagem de que, para
25 gerar um sinal binaural, um mix de canais múltiplos é sempre primeiramente necessário. Ou seja, os sinais de canal múltiplo (por exemplo, canais 5+1) são primeiramente decodificados e sintetizados e as HRTFs são, então, aplicadas a cada sinal para formar um sinal binaural. Esta é, de forma computacional, uma abordagem intensa quando comparada à decodificação direta a partir do formato de canal múltiplo comprimido em
30 formato binaural.

Binaural Cue Coding (BCC) é um método de codificação de áudio espacial paramétrico altamente desenvolvido. A BBC representa um sinal de canal múltiplo especial como um único (ou diversos) canal de áudio com redução de canais e um conjunto de diferenças de intercanal relevante de forma perceptiva estimado como uma
5 função de frequência e tempo a partir do sinal original. O método permite que um sinal de áudio especial seja mixado para um planejamento de alto-falante arbitrário a ser convertido para qualquer outro alto-falante, que consiste no mesmo ou em um número diferente de alto-falantes.

Consequentemente, a BBC é projetada para sistemas de alto-falante de canal
10 múltiplo. Entretanto, gerar um sinal binaural a partir de um sinal mono processado por BBC e suas informações secundárias exige que uma representação de canal múltiplo seja primeiramente sintetizada com base no sinal mono e nas informações secundárias e, somente depois, é possível gerar um sinal binaural para a reprodução espacial em alto-falantes, a partir da representação de canal múltiplo. Fica evidente que tal abordagem não
15 é otimizada em vista de gerar um sinal binaural.

Sumário da Invenção

Atualmente, foi inventado um método e equipamento técnicos aperfeiçoados que executam o método, através do qual é permitida a geração de um sinal binaural a partir de um sinal de áudio parametricamente codificado. Diversos aspectos da invenção
20 incluem um método de decodificação, um decodificador, um aparelho, um método de codificação, um codificador e programas de computadores, os quais são caracterizados pelo o que é afirmado nas reivindicações independentes. Diversas modalidades da invenção são apresentadas nas reivindicações dependentes.

De acordo com um primeiro aspecto, um método, de acordo com a
25 invenção, é baseado na idéia de sintetização de um sinal de áudio binaural, de modo que um sinal de áudio parametricamente codificado que compreende, ao menos, um sinal combinado de uma pluralidade de canais de áudio e um ou mais conjuntos correspondentes das informações secundárias que descreve uma imagem sonora de canal múltiplo é primeiramente inserido. Então, um conjunto pré-determinado filtros de função de
30 transferência relacionados à cabeça é aplicado a, ao menos, um sinal combinado em

proporção determinada pelo dito conjunto de informações secundárias correspondente para sintetizar um sinal de áudio binaural.

De acordo com uma modalidade, a partir do conjunto pré-determinado de filtros de função de transferência relacionados à cabeça, um par esquerdo-direito de filtros de função de transferência relacionados à cabeça correspondente a cada direção de alto-falante do planejamento do alto-falante de canal múltiplo original é escolhido para ser aplicado.

De acordo com uma modalidade, o dito conjunto de informações secundárias compreende um conjunto de estimativas de ganho para os sinais de canal do áudio de canal múltiplo, descrevendo-se a imagem sonora original.

De acordo com uma modalidade, as estimativas de ganho do áudio de canal múltiplo original são determinadas como uma função de tempo e frequência; e os ganhos para cada canal de alto-falante são ajustados de maneira que a soma dos quadrados de cada valor de ganho é igual a um.

De acordo com uma modalidade, ao menos, um sinal combinado é dividido em quadros de tempo de um comprimento de quadro empregado, tais quadros são, então, providos de janelas; e, ao menos, um sinal combinado é transformado no domínio de frequência anterior à aplicação dos filtros de função de transferência relacionados à cabeça.

De acordo com uma modalidade, ao menos, um sinal combinado é dividido no domínio de frequência em uma pluralidade de bandas de frequência motivadas psico-acusticamente, como as bandas de frequência que cumprem com a escala de Largura de Banda Retangular Equivalente (ERB), antes de aplicar os filtros de função de transferência relacionados à cabeça.

De acordo com uma modalidade, as saídas dos filtros de função de transferência relacionados à cabeça para cada banda de frequência para um sinal de lado esquerdo e um sinal de lado direito são somadas separadamente; e o sinal de lado esquerdo somado e o sinal de lado direito somado são transformados no domínio de tempo para criar um componente de lado esquerdo e um componente de lado direito de um sinal de áudio binaural.

Um Segundo aspecto fornece um método para gerar um sinal de áudio parametricamente codificado, o método compreende: inserir um sinal de áudio de canal múltiplo que compreende uma pluralidade de canais de áudio; gerar, ao menos, um sinal combinado da pluralidade de canais de áudio; e gerar um ou mais conjuntos correspondentes de informações secundárias que incluem estimativas de ganho para a pluralidade de canais de áudio.

De acordo com uma modalidade, as estimativas de ganho são calculadas através da comparação do nível de ganho de cada canal individual ao nível de ganho acumulado do sinal combinado.

A disposição, de acordo com a invenção, fornece vantagens significantes. Uma grande vantagem é a simplicidade e a baixa complexidade computacional do processo de decodificação. O decodificador também é flexível no sentido de que realiza a síntese binaural de forma completa com base nos parâmetros espaciais e codificadores dados pelo codificador. Além disso, uma espacialidade equivalente ao sinal original é mantida na conversão. Já para as informações secundárias, um conjunto de estimativas de ganho do mix original é suficiente. De modo mais significativo, a invenção permite uma exploração aperfeiçoada do estado intermediário compressivo fornecido na codificação de áudio paramétrico, aprimorando a eficiência de transmissão, bem como da armazenagem do áudio.

Os aspectos adicionais da invenção incluem diversos aparelhos dispostos para desenvolver as etapas inventivas dos métodos acima.

Breve Descrição dos Desenhos

A seguir, diversas modalidades da invenção serão descritas em maiores detalhes com referência aos desenhos em anexo, nos quais:

A Figura 1 mostra um esquema de Binaural Cue Coding (BCC) genérico, de acordo com uma técnica anterior;

A Figura 2 mostra a estrutura geral de um esquema de síntese de BBC, de acordo com uma técnica anterior;

A Figura 3 mostra um diagrama em bloco do decodificador binaural, de acordo com uma modalidade da invenção; e

A Figura mostra um dispositivo eletrônico, de acordo com uma modalidade da invenção, em um gráfico em bloco reduzido.

Descrição das Modalidades

A seguir, a invenção será ilustrada através de referências ao Binaural Cue Coding (BCC) como uma plataforma exemplificativa para executar o esquema de decodificação, de acordo com as modalidades. Pode-se observar, no entanto, que a invenção não se limita somente aos métodos de codificação de áudio espacial do tipo BBC, porém pode ser executada em qualquer esquema de codificação de áudio que forneça, ao menos, um sinal de áudio combinado a partir do conjunto original de um ou mais canais de áudio e informações secundárias espaciais apropriadas.

Binaural Cue Coding (BCC) é um conceito geral para representação paramétrica de áudio especial, que transfere saída de canal múltiplo com um número arbitrário de canais a partir de um único canal de áudio, além de algumas informações secundárias. A Figura 1 ilustra tal conceito. Diversos canais de áudio de entrada (M) são combinados em um único sinal de saída (S; “soma”) por um processo de redução de canais. Em paralelo, os indicadores de intercanal mais notáveis que descrevem a imagem de som de canal múltiplo são extraídas a partir dos canais de entrada e codificadas de forma compacta como as informações secundárias de BBC. Tanto o sinal de soma quanto a informações secundárias são, então, transmitidos ao lado receptor, possivelmente usando-se um esquema de codificação de áudio de baixa taxa de transferência para codificar o sinal de soma. Finalmente, o decodificador de BBC gera um sinal de saída de (N) de canal múltiplo para alto-falantes a partir do sinal de soma transmitido e a informação de indicador especial através da re-sintetização de canais de saída de canal, os quais podem transportar os indicadores de intercanal relevantes, como Diferença de Tempo de Intercanal (ICTD), Diferença de Nível de Intercanal (ICLD) e Coerência de Intercanal (ICC). Consequentemente, a informações secundárias de BBC, isto é, os indicadores de intercanal, é escolhida em vista da otimização da reconstrução do sinal de áudio de canal múltiplo particularmente para reprodução por alto-falantes.

Há dois esquemas de BBC, a saber, BBC para Renderização Flexível (BBC do tipo I), a qual tem o propósito de transmitir um número de sinais de fonte separados

com o objetivo de renderização no receptor, e BCC para Renderização Natural (BBC do tipo II), a qual tem o propósito para transmissão de um número de canais de áudio de sinal estéreo ou ambiente. A BCC para Renderização Flexível exige sinais de fonte de áudio separados (por exemplo, sinais de fala, instrumentos gravados separadamente, gravação de

5 múltiplas trilhas) como entrada. A BCC para Renderização Natural, por sua vez, exige um sinal de canal múltiplo ou estéreo de “mix final” como entrada (por exemplo, áudio de CD, ambiente de DVD). Caso tais processos sejam realizados através de técnicas de codificação convencionais, as escalas de taxa de transferência proporcionalmente ou, ao menos, de maneira quase proporcional ao número de canais de áudio, por exemplo,

10 transmitir os seis canais de áudio do sistema de canal múltiplo 5.1. exige uma taxa de transferência de aproximadamente seis vezes do canal de áudio. Portanto, ambos os esquemas de BBC resultam em uma taxa de transferência, a qual é somente ligeiramente mais alta do que a taxa de transferência exigida para a transmissão de um canal de áudio, já que a informações secundárias de BBC exige somente uma taxa de transferência muito

15 baixa (e.g. 2 kb/s).

A Figura 2 mostra a estrutura geral de um esquema de síntese de BBC. O sinal mono transmitido (“soma”) é primeiramente provido de janelas no domínio de tempo em quadros e, então, mapeados para uma representação espectral de sub-bandas adequadas por um processo de FFT (Transformação de Fourier Rápida) e um banco de filtros FB.

20 Ao invés dos processos na FFT e FB, um processo de banco de filtro de QMF (Filtro Espelhado em Quadratura) pode ser usado para realizar uma decomposição do sinal. No caso geral de canais de reprodução, a ICLD e a ICTD são consideradas em cada sub-banda entre pares de canais, isto é, para cada canal relativo a um canal de referência. As sub-bandas são selecionadas de maneira que uma resolução de frequência suficientemente

25 alta é alcançada, por exemplo, uma largura de sub-banda igual ao dobro da escala ERB (Largura de Banda Retangular Equivalente) é considerada tipicamente adequada. Para cada canal de saída a ser gerado, a ICTD de atraso de tempo individuais e a ICLD de diferenças de nível são impostas nos coeficientes espectrais, seguidas por um processo de síntese de coerência que reintroduz os aspectos mais relevantes de coerência e/ou

30 correlação (ICC) entre os canais de áudio sintetizados. Finalmente, todos os canais de

saída sintetizados são convertidos em uma representação de domínio de tempo através de um processo de IFFT (FFT Inverso), que resulta na saída de canal múltiplo. Para uma descrição mais detalhada da abordagem de BCC, uma referência é feita a: F. Baumgarte e C. Faller: *"Binaural Cue Coding – Part I: Psychoacoustic Fundamentals and Design Principles"*; IEEE Transactions on Speech and Audio Processing, Vol. 11, Nº 6, novembro de 2003, and to: C. Faller and F. Baumgarte: *"Binaural Cue Coding – Part II: Schemes and Applications"*, IEEE Transactions on Speech and Audio Processing, Vol. 11, No. 6, novembro de 2003.

A BCC é um exemplo de esquemas de codificação, a qual fornece uma plataforma adequada para executar o esquema de decodificação, de acordo com as modalidades. O decodificador binaural, de acordo com uma modalidade, recebe o sinal monofônico e a informações secundárias como entradas. A idéia é substituir cada alto-falante no mix original por um par de HRTFs que correspondem à direção do alto-falante em relação à posição de escuta. Cada canal de frequência do sinal monofônico é alimentado para cada par de filtros que executa as HRTFs na proporção imposta por um conjunto de valores de ganho, os quais podem ser calculados com base na informações secundárias. Consequentemente, o processo pode ser considerado como a execução de um conjunto de alto-falantes virtuais, correspondentes aos originais, na cena de áudio binaural. Desta forma, a invenção adiciona valor à BCC através da permissão para, além de sinais de áudio para canais múltiplos para diversos planejamentos de alto-falante, também um sinal de áudio binaural a ser derivado diretamente a partir de sinal de áudio espacial parametricamente codificado sem nenhum processo de síntese de BCC intermediário.

Algumas modalidades da invenção são ilustradas a seguir com referência à Figura 3, a qual mostra um diagrama em bloco de um decodificador binaural, de acordo com um aspecto da invenção. O decodificador 300 compreende uma primeira entrada 302 para o sinal monofônico e uma segunda entrada 304 para a informações secundárias. As entradas 302, 304 são mostradas como entradas distintivas com o objetivo de ilustrar as modalidades, porém uma pessoa versada na técnica certifica-se de que, em implantação

prática, o sinal monofônico e a informações secundárias podem ser fornecidos através da mesma entrada.

De acordo com uma modalidade, a informações secundárias não tem que incluir os mesmos indicadores de intercanal como nos esquemas de BBC, isto é, Diferença de Tempo de Intercanal (ICTD), Diferença de Nível de Intercanal (ICLD) e Coerência de Intercanal (ICC), porém somente um conjunto de estimativas de ganho que definem a distribuição de pressão de som entre os canais do mix original a cada banda de frequência são suficientes. Além das estimativas de ganho, a informações secundárias inclui, de preferência, o número e locais dos alto-falantes do mix original em relação à posição de escuta, bem como o comprimento do quadro empregado. De acordo com uma modalidade, ao invés de transmitir as estimativas de ganho como uma parte da informações secundárias a partir de um codificador, as estimativas de ganho são computadas no decodificador a partir dos indicadores de intercanal dos esquemas de BBC, por exemplo, a partir da ICLD.

O decodificador 300 compreende, ainda, uma unidade de janelas 306, em que o sinal monofônico é primeiramente dividido em quadros de tempo do comprimento de tempo empregado, e, então, os quadros são divididos em janelas adequadamente, por exemplo, janelas senoidais. Um comprimento de quadro adequado deve ser ajustado de modo que os quadros sejam longos o suficiente para transformação discreta de Fourier (DFT), enquanto é, simultaneamente, curto o suficiente para conduzir rápidas variações no sinal. Experimentos mostraram que o comprimento de quadro adequado é de cerca de 50 ms. Consequentemente, se a frequência de amostra de 44,1 kHz (comumente usada em diversos esquemas de codificação de áudio) é usada, então, o quadro pode compreender, por exemplo, 2048 amostras que resultam no comprimento de quadro de 46,4 ms. A formação das janelas é feita, de preferência, de modo que janelas adjacentes estão sobrepostas por 50% para uniformizar as transições causadas por modificações espectrais (nível e atraso).

Com o propósito de computar de modo eficiente o sinal com domínio em frequência, o sinal é alimentado no banco de filtro 310, o qual divide o sinal em bandas de frequência psico-acusticamente motivadas. De acordo com uma modalidade, o banco de

filtro 310 é projetado de tal forma que está disposto para dividir o sinal em 32 bandas de frequência, de acordo com a escala, comumente conhecida como largura de banda retangular equivalente (ERB), resultando em componentes de sinal $x_0, \dots, x_{31}$ nas ditas 32 bandas de frequência.

5 Como uma alternativa para os blocos 306, 308 e 310, o sinal com domínio tempo-frequência que processa o sinal monofônico pode ser executado em uma unidade QMF de banco-filtro que desempenha a decomposição do sinal. Uma pessoa versada na técnica se certifica que, além do processamento FFT ou um processamento de banco-filtro de QMF, qualquer outro método adequado para executar o processamento do domínio
10 tempo-frequência desejado, pode ser usado.

O decodificador 300 compreende um conjunto de HRTFs 312, 314 como informação pré-armazenada, da qual um par de esquerda-direita de HRTFs correspondente a cada direção do alto-falante, é selecionado. Em consideração à ilustração, dois conjuntos de HRTFs 312, 314 são mostrados na Figura 3, um para o sinal de lado esquerdo e um
15 para o sinal de lado direito, mas é evidente que na implantação prática, um conjunto de HRTFs será suficiente. Para ajustar os pares de HRTFs esquerdo-direita escolhidos para corresponder a cada nível de som do canal de alto-falante, os valores de ganho G são preferivelmente estimados. Como foi mencionado acima, as estimativas de ganho podem ser incluídas na informação secundário recebida a partir do codificador ou podem ser calcula-
20 das no decodificador com base na informação secundário BBC. Consequentemente, um ganho é estimado por cada canal do alto-falante como uma função de tempo e frequência e para preservar o nível de ganho da mix original, os ganhos para cada canal de alto-falante são ajustados preferivelmente de forma que a soma dos quadrados de cada valor dos ganhos seja igual a um, o que fornece a vantagem que, caso N seja o número dos canais a
25 ser virtualmente gerada, então apenas as estimativas de ganho de $N-1$ precisam ser transmitidas a partir de um codificador e a perda de valor de ganho pode ser calculada com base nos valores de ganho $N-1$. Uma pessoa versada na técnica, entretanto, se certifica que a operação da invenção não precisa de ajuste da soma dos quadrados de cada valor de ganho para ser igual a um, mas o codificador pode fazer a escala dos quadrados dos valores de
30 ganho de forma que a soma seja igual a um.

Portanto, cada par esquerdo-direita dos filtros HRTF 312, 314 são ajustados na proporção ditada pelo conjunto dos ganhos G , resultando na adição dos filtros HRTF 312', 314'. Mais uma vez, nota-se que na prática, a escala das magnitudes do filtro original HRTF 312, 314 é meramente determinada, de acordo com os valores de ganho, porém em consideração à ilustração das modalidades, os conjuntos adicionais de HRTFs 312', 314' são mostrados na Figura 3.

Para cada banda de frequência, os componentes de sinal mono $x_0, \dots, x_{31}$ alimentam cada par esquerdo-direita dos filtros HRTF 312', 314' ajustados. As saídas do filtro para o sinal de lado esquerdo e para o sinal de lado direito são, então, somados em unidades de soma 316, 318 para ambos os lados dos canais binaurais. Os sinais binaurais somados são providos com janelas senoidais novamente e transformados mais uma vez no domínio de tempo por um processo inverso FFT executado nas unidades IFFT 320, 322. No caso dos filtros de análise não somarem um ou suas respostas de fase não serem lineares, um banco de filtro de síntese adequada é então, preferivelmente usado para evitar distorção nos sinais finais binaurais B_R e B_L . Mais uma vez, se uma unidade banco-filtro QMF é usada na decomposição do sinal como foi descrito acima, as unidades IFFT 320, 322 são preferencialmente substituídas pelas unidades de banco-filtro (IQMF inverso).

De acordo com uma modalidade, de maneira a intensificar a externalização, isto é, a localização fora da cabeça do sinal binaural e uma resposta de ambiente moderado podem ser adicionadas ao sinal binaural. Com este propósito, o decodificador pode compreender uma unidade de reverberação, localizada, preferivelmente, entre as unidades de soma 316, 318 e as unidades IFFT 320, 322. A resposta de ambiente moderado imita o efeito do ambiente em uma situação de escuta por alto-falante. O tempo de reverberação necessário é, entretanto, curto o bastante para que a complexidade computacional não seja consideravelmente intensificada.

O decodificador binaural 300 descrito na Figura 3 também permite que um caso especial de um decodificador estéreo redução de canais, no qual a imagem espacial é reduzido. A operação do decodificador 300 é emendada de forma que cada filtro HRTF ajustável 312, 314, em que nas modalidades que tiveram suas escalas meramente determinadas, de acordo com os valores de ganho, foram substituídos por um ganho pré-

determinado. Conseqüentemente, o sinal monofônico é processado através de filtros de HRTF constantes que consistem em um único ganho multiplicado pelo conjunto de valores de ganho calculado com base na informação secundária. Como resultado, o áudio especial é misturado a um sinal estéreo. Esse caso especial fornece a vantagem que um sinal estéreo pode ser criado a partir do sinal combinado, usando-se a informação secundária espacial sem que se tenha a necessidade de decodificar o áudio espacial, onde o procedimento da decodificação do estéreo é mais simples que a síntese convencional BCC. A estrutura do decodificador binaural 300 permanece, de qualquer forma, a mesma nas Figura 3, apenas os filtros ajustáveis de HRTF 312, 314 são substituídos por filtros dotados de ganhos pré-determinados para pelo estéreo redução de canais.

Caso o decodificador binaural compreender filtros de HRTF, por exemplo, por uma configuração de áudio ambiente 5.1, e então no caso espacial para decodificação da redução de canais estéreo, os ganhos constantes dos filtros de HRTF podem ser, por exemplo, como definidos na Tabela 1.

HRTF	Esquerda	Direita
Frontal do lado esquerdo	1,0	0,0
Frontal do lado direito	0,0	1,0
Centro	raiz quadrada (0,5)	raiz quadrada (0,5)
Posterior do lado esquerdo	raiz quadrada (0,5)	0,0
Posterior do lado direito	0,0	raiz quadrada (0,5)
LFE	raiz quadrada (0,5)	raiz quadrada (0,5)

Tabela 1, filtros de HRTF para estéreo redução de canais.

A disposição de acordo com a invenção fornece vantagem. Uma vantagem ainda maior é a simplicidade e a baixa complexidade computacional do processo de decodificação. O decodificador também é flexível no que diz respeito ao desempenho completo do aumento de canais binaural com base nos parâmetros espaciais e de decodificação. Além disso, com referência ao espaço, o sinal original é mantido na conversão. Quanto à informação secundária, um conjunto da estimativa de ganho do mix original é suficiente. Do ponto de vista da transmissão ou armazenamento de áudio, a vantagem mais signifi-

cante é obtida através da eficiência aprimorada ao se utilizar o estado compressivo intermediário fornecido na codificação de áudio paramétrica.

Um profissional versado na técnica se certifica que, desde que os HRTFs sejam altamente individuais e obter a média é impossível, a re-espacialização pode apenas ser alcançada pela medição do único conjunto HRTFs do próprio ouvinte. Consequentemente, o uso de HRTFs inevitavelmente coloriza o sinal de forma que a qualidade do áudio processado não é equivalente a do original. De qualquer forma, uma vez que a medição de cada HRTFs dos ouvintes seja uma opção inviável, o melhor resultado possível será alcançado, quando tanto os conjuntos moldados quanto um conjunto medido por uma cabeça artificial ou uma pessoa com uma cabeça de tamanho médio e de notável simetria, será usado.

Como foi determinado anteriormente, de acordo com uma modalidade, as estimativas de ganho podem ser incluídas na informação secundária recebida a partir do codificador. Consequentemente, um aspecto da invenção refere-se a um codificador para sinal de áudio especial de canal múltiplo que estima um ganho para cada canal de alto-falante como uma função de frequência e tempo e inclui as estimativas de ganho na informação secundária a serem transmitidas junto com um (ou mais) canal(is) combinado(s). O codificador pode ser, por exemplo, um codificador de BCC conhecido como tal, que está disposto mais adiante, para calcular as estimativas de ganho, ambos que em adição e ao invés disso, os indicadores de intercanal ICTD, ICLD e ICC descrevendo a imagem do som de canal múltiplo. Portanto, ambas as somas do sinal e da informação secundária, que compreendem, ao menos, as estimativas de ganho, são transmitidas ao lado do receptor, usando-se preferivelmente um esquema adequado de codificação de áudio de baixa taxa de transferência adequada, transmitido ao lado receptor, usando preferencialmente um esquema adequado de codificação de áudio de baixa taxa de transferência para codificar o sinal de soma.

De acordo com uma modalidade, se as estimativas de ganho são calculadas no codificador, o cálculo é executado por meio de uma comparação do nível de ganho de cada canal individual com o nível de ganho acumulado do canal combinado; isto é, se denominarmos os níveis de ganho como X , os canais individuais do projeto do alto falante

original como “m” e exemplos como “k” e então, para cada canal, a estimativa de ganho é calculada como $|X_m(k)| / |X_{SOMA}(k)|$. Consequentemente, a estimativa de ganho determina a magnitude do ganho proporcional de cada canal individual em comparação à magnitude do ganho total de todos os canais.

5 De acordo com uma modalidade, se as estimativas de ganho são calculadas no codificador com base da informação secundário BCC, o cálculo pode ser efetuado, por exemplo, na base dos valores do Diferença de Níveis do Canal Interno ICLD. Consequentemente, se N é o número de “alto-falantes” a serem virtualmente gerados, então as equações N-1, que compreende N-1 desconhecidas variam, são primeiramente compostas com
10 base dos valores ICLD. Portanto, a soma dos quadrados de cada equação dos alto-falantes é igual a 1, de forma que a estimativa de ganho de um canal individual pode ser resolvida e com base da estimativa de ganho resolvida, o resto das estimativas de ganho podem ser resolvidas a partir das equações N-1.

Por exemplo, se o número de canais gerado virtualmente for 5, (N=5), as
15 equações N-1 podem ser formadas como a seguir: $L2=L1+ICLD1$, $L3=L1+ICLD2$, $L4=L1+ICLD3$ e $L5=L1+ICLD4$. Daí em diante, a soma dos quadrados é igual a 1: $L1^2 + (L1+ICLD1)^2 + (L1+ICLD2)^2 + (L1+ICLD3)^2 + (L1+ICLD4)^2 = 1$. O valor de L1 pode então, ser resolvido e com base em L1, o resto dos valores dos níveis de ganho L2 – L5 podem ser resolvidos.

20 De modo a simplificar, os exemplos anteriores são descritos de tal forma que os canais de entrada (M) são submetidos à redução de canais no codificador para formar um único canal combinado (por exemplo, mono). De qualquer forma, as modalidades são igualmente aplicáveis em implantações alternativas, onde os múltiplos canais de entrada (M) são reduzidos para formar dois ou mais canais separados (S), dependendo da aplicação particular de processamento de áudio. Se a redução de canais gera múltiplos canais
25 combinados, os dados de canal combinados podem ser transmitidos usando técnicas convencionais de transmissão de áudio. Por exemplo, se dois canais combinados são gerados, técnicas convencionais de transmissão estéreo serão empregadas. Nesse caso, um decodificador BCC pode extrair e usar os códigos BCC para sintetizar um sinal binaural a partir
30 dos dois canais combinados.

De acordo com uma modalidade, o número (N) dos “alto-falantes” gerados virtualmente no sinal binaural sintetizado podem ser diferentes (maiores ou menores) que o número dos canais de entrada (M), dependendo da aplicação particular. Por exemplo, o áudio de entrada pode corresponder a 7,1 som ambiente e a saída binaural de áudio pode ser sintetizada para corresponder a 5,1 de som ambiente ou vice versa.

As modalidades acima podem ser generalizadas de tal forma que as modalidades da invenção permitem a conversão M de canais de áudio em canais de áudio combinados S e um ou mais conjuntos correspondentes de informação secundário, onde $M > S$, e a geração de canais de saída de áudio e os conjuntos correspondentes de informações laterais, onde $N > S$, e N pode ser igual ou diferente de M.

Já que a taxa de transferência é requerida para a transmissão do canal combinado e a informação secundária necessária é muito lenta, a invenção é especialmente bem aplicável nos sistemas, onde a largura de banda disponível é um recurso escasso, tanto quanto em sistemas de comunicação. Consequentemente, as modalidades são especialmente disponíveis em terminais móveis ou em outros dispositivos portáteis são especialmente aplicáveis em terminais móveis ou em outros dispositivos portáteis tipicamente desprovido dos alto-falantes de alta qualidade, onde as características de som ambiente de canais múltiplos podem ser introduzidos através de fones de ouvido, ouvindo o sinal binaural de áudio de acordo com as modalidades. Um campo adicional de aplicações viáveis inclui serviços de teleconferência, em que os participantes de teleconferência podem ser facilmente diferenciados, dando a impressão aos ouvintes, que os participantes da chamada de teleconferência estão em diferentes locais da sala de conferências.

A Figura 4 ilustra uma estrutura simplificada do dispositivo de processamento de dados (TE), onde o sistema binaural de decodificação, de acordo com a invenção pode ser implantada. O dispositivo de processamento (TE) pode ser, por exemplo, um terminal móvel, um dispositivo PDA ou um computador pessoal (PC). A unidade de processamento de dados (TE) compreende meios I/O (I/O), uma unidade central de processamento (CPU) e memória (MEM). A memória (MEM) compreende, a porção ROM de memória somente leitura e uma porção regravável, tal qual uma memória de acesso aleatório RAM e memória FLASH. A informação usada para comunicar-se com partes exter-

nas diferentes, por exemplo, um CD-ROM, outros dispositivos e o usuário, é transmitido através dos meios I/O (I/O) a para/a partir de, a unidade de processamento central (CPU). Se o dispositivo de processamento de dados for implantado como uma estação móvel, isso inclui tipicamente um transceptor Tx/Rx, que se comunica com uma rede sem fio, tipicamente com uma estação de transceptor base (BTS) através de uma antena. O equipamento da Interface do usuário (UI) normalmente inclui um visor, um teclado, um microfone e meios de conexão para fones de ouvido. O dispositivo de processamento de dados pode compreender adicionalmente, meios de conexão MMC, tal qual uma entrada de forma padrão, para vários módulos de hardware ou como circuitos integrados IC, que podem fornecer várias aplicações a serem executadas no dispositivo de processamento de dados.

Consequentemente, o sistema de decodificação binaural de acordo com a invenção, pode ser executado em uma unidade central de processamento CPU ou em um processador de sinais digitais exclusivo DSP (um processador de código paramétrico) do dispositivo de processamento de dados em que o dispositivo de processamento de dados recebe um sinal de áudio parametricamente codificado que compreende ao menos um sinal combinado de vários canais de áudio e um ou mais conjuntos de informações laterais correspondentes, descrevendo uma imagem sinal de áudio de canal múltiplo. O sinal de áudio parametricamente codificado pode ser recebido a partir de um meio de memória, por exemplo, um CD-ROM ou a partir de uma rede sem fio através da antena e o transceptor Tx/Rx. O dispositivo de processamento de dados, compreende adicionalmente um banco de filtro adequado e um conjunto de filtros de função de transferência relacionados à cabeça pré-determinados, em que o dispositivo de processamento de dados transforma o sinal combinado no domínio da frequência e aplica uns pares esquerdo-direita adequados filtros de função de transferência relacionados à cabeça em proporção de sinal combinado determinado pelo conjunto correspondente da informação secundário para sintetizar um sinal de áudio binaural, que é então reproduzido via fones de ouvido.

Da mesma forma, o sistema de codificação de acordo com a invenção pode também ser executada na unidade central de processamento ou em um processador de sinais digitais exclusivo DSP do dispositivo de processamento de dados, em que o dispositivo de processamento de dados gera um sinal de áudio codificado parametricamente, com-

preendendo, ao menos, um sinal combinado de uma pluralidade de canais de áudio e ou mais conjuntos de informações laterais incluindo estimativas de ganho para os sinais de canal do áudio de canal múltiplo.

5 As funcionalidades da invenção podem ser implantadas em um dispositivo terminal, tal qual uma estação móvel, também como um programa de computador que, quando executado em uma unidade de processamento central CPU ou em um processador de sinais digitais DSP, afeta o dispositivo terminal para implantar os procedimentos da invenção. As funções do programa de computador SW podem ser distribuídos em vários componentes de programas separados, comunicando-se entre si. O software do computa-
10 dor pode ser armazenado em quaisquer meios de memória, como por exemplo, o disco rígido de um PC ou um disco de CD-ROM, do qual, ele pode ser carregado na memória do terminal móvel. O software do computador pode também ser carregado através de uma rede, por exemplo, usando uma pilha de protocolo TCP/IP.

Também é possível que se use as soluções de hardware ou uma combinação
15 de soluções de hardware e software para implantar os meios inventivos. Consequentemente, o produto de computador pode ser, ao menos parcialmente implantados como uma solução de hardware, por exemplo como os circuito ASIC ou FPGA, em um módulo de hardware compreendendo os meios de conexão para conectar um dispositivo de módulo a um dispositivo eletrônico ou como um ou mais circuitos integrados IC, o módulo de
20 hardware ou os ICs incluem adicionalmente vários meios para desempenhar as tarefas de código de programa, os ditos meios sendo implantados como hardware ou software..

Fica evidente que a presente invenção não se limita somente às modalidades apresentadas acima, porém elas podem ser modificadas com o escopo das reivindicações anexadas.

REIVINDICAÇÕES

1. Método para a sintetização de um sinal de áudio binaural, sendo que o método é **caracterizado** pelo fato de que compreende:

inserir um sinal de áudio parametricamente codificado que compreende, ao menos, um sinal combinado de uma pluralidade de canais de áudio e um ou mais conjuntos de informações secundárias correspondentes que descrevem uma imagem sonora de canal múltiplo, e

aplicar um conjunto pré-determinado de filtros de função de transferência relacionados à cabeça a, ao menos, um sinal combinado em proporção determinada pelo mencionado conjunto de informações secundárias correspondente para sintetizar um sinal de áudio binaural.

2. Método, de acordo com a reivindicação 1, **caracterizado** pelo fato de que compreende adicionalmente:

aplicar, a partir do conjunto pré-determinado de filtros de função de transferência relacionados à cabeça, um par esquerdo-direito de filtros de função de transferência relacionados à cabeça que correspondem a cada direção do alto-falante do áudio de canal múltiplo original.

3. Método, de acordo com a reivindicação 1 ou 2, **caracterizado** pelo fato de que:

o dito conjunto de informações secundárias compreende um conjunto de estimativas de ganho para os sinais de canal do áudio de canal múltiplo que descrevem a imagem de som original.

4. Método, de acordo com a reivindicação 3, **caracterizado** pelo fato de que:

o dito conjunto de informações secundárias compreende, adicionalmente, o número e locais de alto-falantes da imagem de som de canal múltiplo original em relação a uma posição de escuta e um comprimento de quadro empregado.

5. Método, de acordo com a reivindicação 1 ou 2, **caracterizado** pelo fato de que:

o dito conjunto de informações secundárias compreende indicadores de in-

tercanal usados em um esquema de Binaural Cue Coding (BCC), como Diferença de Tempo de Intercanal (ICTD), Diferença de Nível de Intercanal (ICLD) e Coerência de Intercanal (ICC), o método compreende adicionalmente:

5 calcular um conjunto de estimativas de ganho do áudio de canal múltiplo original com base em, ao menos, um dos ditos indicadores de intercanal do esquema de BCC.

6. Método, de acordo com quaisquer das reivindicações de 3 a 5, **caracterizado** pelo fato de que compreende adicionalmente:

10 determinar o conjunto de estimativas de ganho do áudio de canal múltiplo original como uma função de tempo e frequência, e

ajustar os ganhos para cada canal de alto-falante de modo que a soma dos quadrados de cada valor de ganho seja igual a um.

7. Método, de acordo com quaisquer das reivindicações precedentes, **caracterizado** pelo fato de que compreende adicionalmente:

15 dividir, ao menos, um sinal combinado em quadros de tempo de um comprimento de quadro empregado, tais quadros são, então, submetidos à formação de janelas; e

transformar, ao menos, um sinal combinado em domínio de frequência anterior à aplicação dos filtros de função de transferência relacionados à cabeça.

20 8. Método, de acordo com a reivindicação 7, **caracterizado** pelo fato de que compreende adicionalmente:

dividir, ao menos, um sinal combinado no domínio de frequência em uma pluralidade de bandas de frequência psico-acusticamente motivadas anterior à aplicação dos filtros de função de transferência relacionados à cabeça.

25 9. Método, de acordo com a reivindicação 8, **caracterizado** pelo fato de que compreende adicionalmente:

dividir, ao menos, um sinal combinado no domínio de frequência em 32 bandas de frequência que estão de acordo com a escala de Largura de Banda Retangular Equivalente (ERB).

30 10. Método, de acordo com quaisquer das reivindicações de 7 a 9, **caracte-**

rizado pelo fato de que:

a etapa de transformação de, ao menos, um sinal combinado no domínio de frequência é realizada usando-se filtros QMF para decompor , ao menos, um sinal combinado.

5 11. Método de acordo com quaisquer das reivindicações de 8 a 10, **caracterizado** pelo fato de que compreende adicionalmente:

somar saídas dos filtros de função de transferência relacionados à cabeça para cada uma das ditas bandas de frequência para um sinal de lado esquerdo e um sinal de lado direito separadamente; e

10 transformar o sinal de lado esquerdo somado e o sinal de lado direito somado no domínio de tempo para criar um componente de lado esquerdo e um componente de lado direito de um sinal de áudio binaural.

12. Método para sintetização de um sinal de áudio estéreo, sendo que o método é **caracterizado** pelo fato de compreender:

15 inserir um sinal de audio parametricamente codificado que compreende, ao menos, um sinal combinado de uma pluralidade de canais de áudio e um ou mais conjuntos de informações secundárias correspondente que descrevem uma imagem sonora de canal múltiplo; e

20 aplicar um conjunto de filtros de redução de canais dotados de valores de ganho pré-determinados a, ao menos, um sinal combinado em proporção determinada pelo dito conjunto de informações secundárias correspondente para sintetizar um sinal de áudio estéreo.

13. Decodificador de áudio paramétrico, **caracterizado** pelo fato de que compreende:

25 um processador de código paramétrico para processamento de um sinal de áudio parametricamente codificado que compreende, ao menos, um sinal combinado de uma pluralidade de canais de áudio e um ou mais conjuntos de informações secundárias correspondentes que descrevem uma imagem sonora de canal múltiplo; e

30 um sintetizador para aplicar um conjunto de filtros de função de transferência relacionados à cabeça pré-determinados a, ao menos, i, sinal combinado em proporção

determinada pelo dito conjunto de informações secundárias correspondente para sintetizar um sinal de áudio binaural.

14. Decodificador, de acordo com a reivindicação 13, **caracterizado** pelo fato de que:

5 o dito sintetizador é disposto para aplicar, a partir do conjunto de filtros de função de transferência relacionados à cabeça pré-determinados, um par esquerdo-direito de filtros de função de transferência relacionados à cabeça correspondentes a cada direção de alto-falante do áudio de canal múltiplo original.

10 15. Decodificador, de acordo com a reivindicação 13 ou 14, **caracterizado** pelo fato de que:

o dito conjunto de informações secundárias compreende um conjunto de estimativas de ganho para sinais de canal do áudio de canais múltiplos que descrevem a imagem sonora original.

15 16. Decodificador, de acordo com a reivindicação de 13 ou 14, **caracteri-**
zado pelo fato de que:

o dito conjunto de informações secundárias compreende indicadores de intercanal usados em um esquema de Binaural Cue Coding (BCC), como Diferença de Tempo de Intercanal (ICTD), Diferença de Nível de Intercanal (ICLD) e Coerência de Inter- Canal (ICC), o decodificador é organizado para:

20 calcular um conjunto de estimativas de ganho do áudio de canal múltiplo original com base em, ao menos, tais indicadores de intercanal do esquema de BCC.

17. Decodificador, de acordo com quaisquer das reivindicações de 13 a 16, **caracterizado** pelo fato de que compreende adicionalmente:

25 um meio para dividir, ao menos, um sinal combinado em quadros de tempo de um comprimento de quadro empregado,

um meio para formar as janelas dos quadros; e

um meio para transformar, ao menos, um sinal combinado no domínio de frequência anterior à aplicação dos filtros de função de transferência relacionados à cabeça.

30 18. Decodificador, de acordo com a reivindicação 17, **caracterizado** pelo

fato de que compreende adicionalmente:

um meio para dividir, ao menos, um sinal combinado no domínio de frequência em uma pluralidade de bandas de frequência motivadas psico-acusticamente anterior à aplicação dos filtros de função de transferência relacionados à cabeça.

5 19. Decodificador, de acordo com a reivindicação 18, **caracterizado** pelo fato de que:

o dito meio para dividir, ao menos, um sinal combinado na frequência de domínio que compreende um banco de filtro disposto para dividir, ao menos, um sinal combinado em 32 bandas de frequência que estão de acordo com a escala de Largura de Banda Retangular Equivalente (ERB).

10

20. Decodificador, de acordo com as reivindicações de 17 a 19, **caracterizado** pelo fato de que:

o meio para transformar, ao menos, um sinal combinado no domínio de frequência compreende filtros de QMF dispostos para decompor, ao menos, um sinal combinado.

15

21. Decodificador, de acordo com quaisquer das reivindicações precedentes de 17 a 20, **caracterizado** pelo fato de que compreende adicionalmente:

uma unidade de soma para somar saídas dos filtros de função de transferência relacionados à cabeça para cada uma das ditas bandas de frequência para um sinal de lado esquerdo e um sinal de lado direito separadamente; e

20

uma unidade de transformação para transformar o sinal de lado esquerdo somado e o sinal de lado direito somado em domínio de tempo para criar um componente de lado direito de um sinal de áudio binaural.

22. Decodificador de áudio paramétrico, **caracterizado** pelo fato de que compreende:

25

um processador de código paramétrico para processamento de um sinal de áudio parametricamente codificado que compreende, ao menos, um sinal combinado de uma pluralidade de canais de áudio e um ou mais conjuntos de informações secundárias correspondentes que descrevem uma imagem sonora de canal múltiplo; e

30

um sintetizador para aplicar um conjunto de filtros de redução de canais do-

tados de valores de ganho pré-determinados para, ao menos, um sinal combinado em proporção determinada pelo dito conjunto de informação correspondente para sintetizar um sinal de áudio estéreo.

23. Produto de programa de computador, armazenado em uma mídia legível por computador e executável em um dispositivo de processamento de dados, para processar um sinal de áudio parametricamente codificado que compreende, ao menos, um sinal combinado de uma pluralidade de canais de áudio e um ou mais conjuntos de informação correspondentes que descrevem uma imagem sonora de canal múltiplo, sendo que o produto de programa de computador é **caracterizado** pelo fato de que compreende:

uma sessão de código de programa de computador para controlar a transformação de, ao menos, um sinal combinado no domínio de frequência; e

uma sessão de código de programa de computador para aplicar um conjunto de filtros de função de transferência relacionados à cabeça pré-determinados para, ao menos, um sinal combinado em proporção determinada pelo dito conjunto de informações secundárias correspondente para sintetizar um sinal de áudio binaural.

24. Aparelho para a sintetização de um sinal de áudio binaural, sendo que o aparelho é **caracterizado** pelo fato de compreender:

um meio para inserir um sinal de áudio parametricamente codificado que compreende, ao menos, um sinal combinado de uma pluralidade de canais de áudio e um ou mais conjuntos de informações secundárias correspondentes que descrevem uma imagem sonora de canal múltiplo;

um meio para aplicar um conjunto pré-determinado de filtros de função de transferência relacionados à cabeça para, ao menos, um sinal combinado em proporção determinada pelo dito conjunto de informações secundárias correspondente para sintetizar um sinal de áudio binaural; e

um meio para fornecer o sinal de áudio binaural no meio de reprodução de áudio.

25. Aparelho, de acordo com a reivindicação 24, **caracterizado** pelo fato de que o dito aparelho é um terminal móvel, um dispositivo de PDA ou um computador pessoal.

26. Método para gerar um sinal de áudio parametricamente codificado, sendo que o método é **caracterizado** pelo fato de que compreende:

inserir um sinal de áudio de canal múltiplo que compreende uma pluralidade de canais de áudio;

5 gerar, ao menos, um sinal combinado da pluralidade de canais de áudio; e
gerar um ou mais conjuntos de informações secundárias correspondentes que incluem estimativas de ganho para a pluralidade de canais de áudio.

27. Método, de acordo com a reivindicação 26, **caracterizado** pelo fato de que compreende adicionalmente:

10 calcular as estimativas de ganho comparando-se o nível de ganho de cada canal individual para o nível de ganho acumulado do sinal combinado.

28. Método, de acordo com a reivindicação 26 ou 27, **caracterizado** pelo fato de que:

15 o dito conjunto de informações secundárias compreende, ainda, o número e locais de alto-falantes de uma imagem sonora de canal múltiplo original em relação à posição de escuta e um comprimento de quadro empregado.

29. Método, de acordo com quaisquer das reivindicações de 26 a 28, **caracterizado** pelo fato de que:

20 o dito conjunto de informações secundárias compreende, ainda, indicadores de intercanal usados em um esquema de Binaural Cue Coding (BCC), como Diferença de Tempo de Intercanal (ICTD), Diferença de Nível de Intercanal (ICLD) e Coerência de Intercanal (ICC).

30. Método, de acordo com quaisquer das reivindicações de 26 a 29, **caracterizado** pelo fato de que compreende adicionalmente:

25 determinar o conjunto de estimativas de ganho do áudio de canal múltiplo original como uma função de tempo e frequência; e

ajustar os ganhos para cada canal de alto-falante de modo que a soma dos quadrados de cada valor de ganho seja igual a um.

30 31. Codificador de áudio paramétrico para gerar um sinal de áudio parametricamente codificado, sendo que o codificador é **caracterizado** pelo fato de compreender:

um meio para inserir um sinal de áudio de canal múltiplo que compreende uma pluralidade de canais de áudio;

um meio para gerar, ao menos, um sinal combinado da pluralidade de canais de áudio; e

5 um meio para gerar um ou mais conjuntos de informações secundárias correspondentes que incluem estimativas de ganho para a pluralidade de canais de áudio.

32. Codificador, de acordo com a reivindicação 31, **caracterizado** pelo fato de que compreende adicionalmente:

10 um meio para calcular as estimativas de ganho comparando-se o nível de ganho acumulado do sinal combinado.

33. Produto de programa de computador, armazenado em uma mídia legível por computador e executável em um dispositivo de processamento de dados, para gerar um sinal de áudio parametricamente codificado, sendo que o produto de programa de computador é **caracterizado** pelo fato de compreender:

15 uma sessão de código de programa de computador para inserir um sinal de áudio de canal múltiplo que compreende uma pluralidade de canais de áudio;

uma sessão de código de programa de computador para gerar, ao menos, um sinal combinado da pluralidade de canais de áudio; e

20 uma sessão de código de programa de computador para gerar um ou mais conjuntos de informações secundárias correspondentes que incluem estimativas de ganho para a pluralidade de canais de áudio.

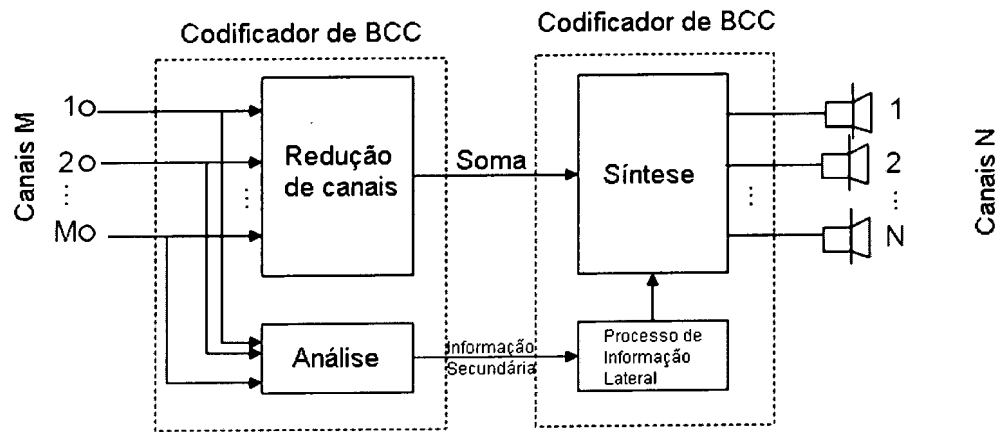


Fig. 1

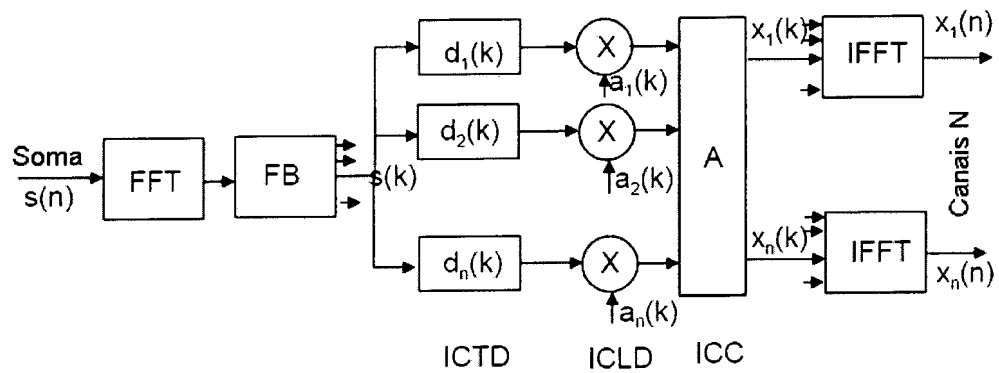


Fig. 2

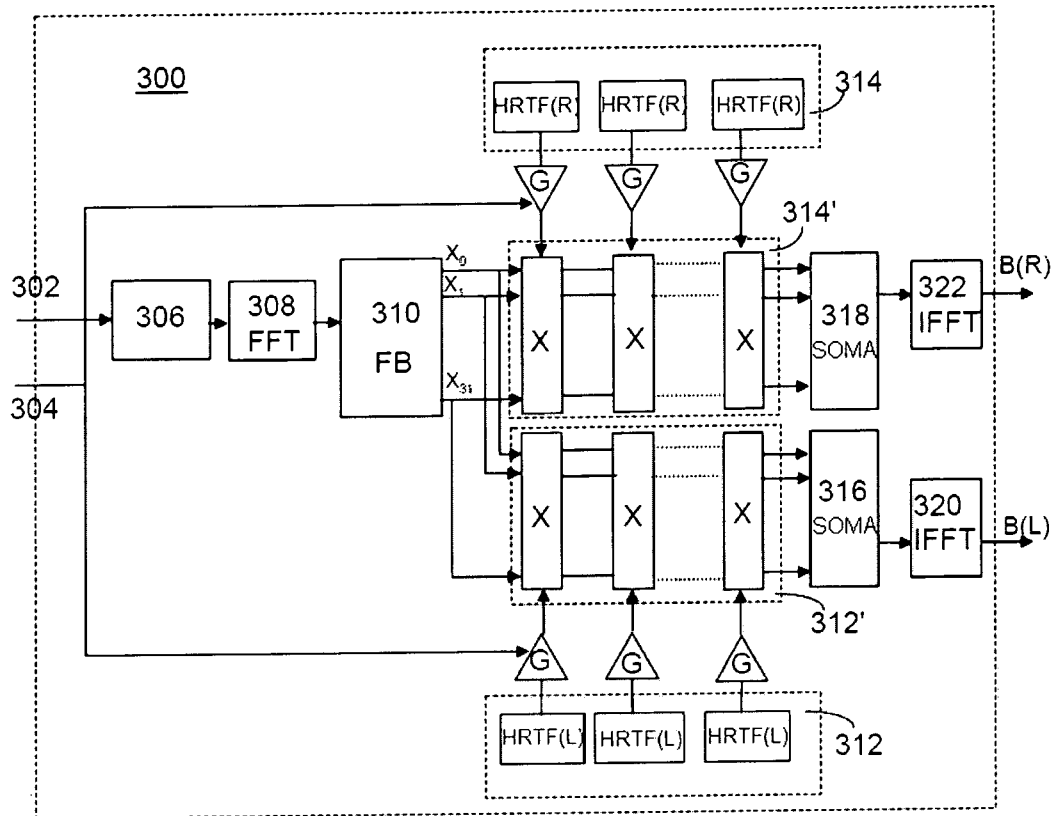


Fig. 3

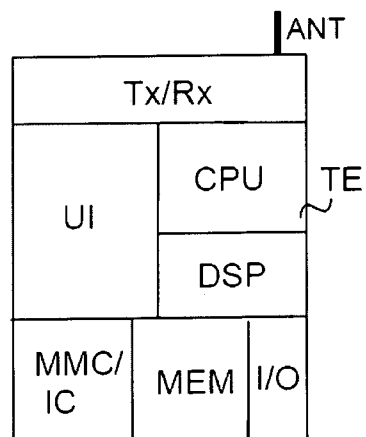


Fig. 4

RESUMO

"MÉTODO E APARELHO PARA A SINTETIZAÇÃO DE UM SINAL DE ÁUDIO BINAURAL; . MÉTODO PARA SINTETIZAÇÃO DE UM SINAL DE ÁUDIO ESTÉREO; DECODIFICADOR DE ÁUDIO PARAMÉTRICO; PRODUTO DE PROGRAMA DE COMPUTADOR, ARMAZENADO EM UMA MÍDIA LEGÍVEL POR COMPUTADOR E EXECUTÁVEL EM UM DISPOSITIVO DE PROCESSAMENTO DE DADOS, PARA PROCESSAR UM SINAL DE ÁUDIO PARAMETRICAMENTE CODIFICADO QUE COMPREENDE, AO MENOS, UM SINAL COMBINADO DE UMA PLURALIDADE DE CANAIS DE ÁUDIO E UM OU MAIS CONJUNTOS DE INFORMAÇÃO CORRESPONDENTES QUE DESCREVEM UMA IMAGEM SONORA DE CANAL MÚLTIPLO; MÉTODO PARA GERAR UM SINAL DE ÁUDIO PARAMETRICAMENTE CODIFICADO; CODIFICADOR DE ÁUDIO PARAMÉTRICO PARA GERAR UM SINAL DE ÁUDIO PARAMETRICAMENTE CODIFICADO; PRODUTO DE PROGRAMA DE COMPUTADOR, ARMAZENADO EM UMA MÍDIA LEGÍVEL POR COMPUTADOR E EXECUTÁVEL EM UM DISPOSITIVO DE PROCESSAMENTO DE DADOS, PARA GERAR UM SINAL DE ÁUDIO PARAMETRICAMENTE CODIFICADO".

Trata-se de um método para sintetizar um sinal de áudio binaural, sendo que o método compreende: inserir um sinal de áudio parametricamente codificado em, ao menos, um sinal combinado de uma pluralidade de canais de áudio e um ou mais conjuntos de informações secundárias que descrevem uma imagem sonora de canal múltiplo; e aplicar um conjunto pré-determinado de filtros de função de transferência relacionados à cabeça para, ao menos, um sinal combinado em proporção determinada pelo conjunto de informações secundárias correspondentes para sintetizar um sinal de áudio binaural. São descritos, ainda, um decodificador de áudio paramétrico, um codificador de áudio paramétrico, um produto de programa de computador e um aparelho para sintetização de um sinal de áudio binaural.