



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

(11) PI 0518507-6 B1



(22) Data do Depósito: 30/09/2005

(45) Data de Concessão: 22/12/2020

(54) Título: INFORMAÇÕES AUXILIARES COMPACTAS PARA A CODIFICAÇÃO PARAMÉTRICA DE ÁUDIO ESPACIAL

(51) Int.Cl.: G10L 19/008; H04S 3/00.

(52) CPC: G10L 19/008; H04S 3/00.

(30) Prioridade Unionista: 10/01/2005 US 11/032,689.

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

(72) Inventor(es): JÜRGEN, HERRE; CHRISTOF, FALLER.

(86) Pedido PCT: PCT EP2005010595 de 30/09/2005

(87) Publicação PCT: WO 2006/072270 de 13/07/2006

(85) Data do Início da Fase Nacional: 10/07/2007

(57) Resumo: INFORMAÇÕES AUXILIARES COMPACTAS PARA A CODIFICAÇÃO PARAMÉTRICA DE ÁUDIO ESPACIAL. Em um codificador de áudio, são gerados códigos de sinalização para um ou mais canais de áudio, caracterizados pelo fato de que um código de sinalização combinado (ex., um código de correlação intercanais (ICC) combinado) é gerado pela combinação de dois ou mais códigos de sinalização estimados, cada código de sinalização estimado sendo estimado a partir de um grupo de dois ou mais canais. Em um decodificador de áudio, E canal(is) de áudio transmitido(s) são decodificados para gerar C canais de áudio playback. Os códigos de sinalização recebidos incluem um código de sinalização combinado (ex., um código ICC combinado) . Um ou mais canal(is) transmitido(s) são upmixados para gerar um ou mais canais upmixados. Um ou mais canais playback são sintetizados pela aplicação dos códigos de sinalização a um ou mais canais upmixados, em que dois ou mais códigos de sinalização derivados se derivam de um código de sinalização combinado, e cada código de sinalização derivado é aplicado para gerar dois ou mais canais sintetizados.

INFORMAÇÕES AUXILIARES COMPACTAS PARA A
CODIFICAÇÃO PARAMÉTRICA DE ÁUDIO ESPACIAL

HISTÓRICO DA INVENÇÃO

Referência Cruzada com os Pedidos Relacionados

[001] O objeto em questão do presente pedido está relacionado com o objeto em questão dos seguintes pedidos norte-americanos, cujos ensinamentos seguem incorporados à presente como referência:

[002] Pedido Norte-Americano Número de Série 09/848,877, depositado em 4 de maio de 2001 como protocolo do agente número Faller 5;

[003] Pedido Norte-Americano Número de Série 10/045,458, depositado em 7 de novembro de 2001 como protocolo do agente número Baumgarte 1-6-8, que reivindica para si o benefício da data de depósito do pedido provisório norte-americano número 60/311,565, depositado em 10 de agosto de 2001;

[004] Pedido Norte-Americano Número de Série 10/155.437, depositado em 24 de maio de 2002 como protocolo do agente número Baumgarte 2-10;

[005] Pedido Norte-Americano Número de Série 10/246.570, depositado em 18 de setembro de 2002 como protocolo do agente número Baumgarte 3-11;

[006] Pedido Norte-Americano Número de Série 10/815.591, depositado em 1º de abril de 2004 como protocolo do agente número Baumgarte 7-12;

[007] Pedido Norte-Americano Número de Série 10/936.464, depositado em 8 de setembro de 2004 como protocolo do agente número Baumgarte 8-7-15;

[0008] Pedido Norte-Americano Número de Série 10/762.100, depositado em 20 de janeiro de 2004 (Faller 13-1);

[0009] Pedido Norte-Americano Número de Série 11/006.492, depositado em 7 de dezembro de 2004 como protocolo do agente número Allamanche 1-2-17-3; e

[0010] Pedido Norte-Americano Número de Série 11/006,_____, depositado em 7 de dezembro de 2004 como protocolo do agente número Allamanche 2-3-18-4.

[0011] O objeto em questão do presente pedido também está relacionado ao objeto em questão descrito nos seguintes documentos, cujos ensinamentos seguem incorporados à presente por referência:

[0012] F. Baumgarte and C. Faller, "Binaural Cue Coding-Part I: Psychoacoustic fundamentals and design principles," IEEE Trans. on Speech and Audio Proc., vol. 11, no. 6, November 2003;

[0013] C. Faller and F. Baumgarte, "Binaural Cue Coding-Part II: Schemes and applications," IEEE Trans. on Speech and Audio Proc., vol. 11, no. 6, November 2003; e

[0014] C. Faller, "Coding of spatial audio compatible with different playback formats," Preprint 117th Conv. Aud. Eng. Soc., October 2004.

Campo da Invenção

[0015] A presente invenção se refere à codificação de sinais de áudio e a subsequente síntese de **cenais auditivas** dos dados de áudio codificados.

Descrição da Técnica Relacionada

[0016] Quando uma pessoa ouve um sinal de áudio (isto é, sons) gerado por uma determinada fonte de áudio, o sinal de áudio tipicamente chegará aos ouvidos direito e esquerdo da pessoa em dois

tempos diferentes e com dois níveis diferentes de áudio (ex., decibéis), sendo que esses diferentes tempos e níveis são funções das diferenças dos caminhos que percorrem os sinais de áudio até atingirem os ouvidos direito e esquerdo, respectivamente. O cérebro da pessoa interpreta essas diferenças em tempo e nível de maneira a dar à pessoa a percepção de que o sinal de áudio recebido está sendo gerado por uma fonte de áudio localizada em uma determinada posição (ex., direção e distância) com relação à pessoa. Uma cena **auditiva** é o efeito líquido de uma pessoa que ouve simultaneamente sinais de áudio gerados por uma ou mais diferentes fontes de áudio localizadas em uma ou mais diferentes posições relativas à pessoa.

[0017] A existência desse processamento pelo cérebro pode ser usada para sintetizar **cenais auditivas**, em que os sinais de áudio provenientes de uma ou mais diferentes fontes de áudio são modificados de propósito para gerar sinais de áudio esquerdo e direito que dão a percepção de que diferentes fontes de áudio estão localizadas em diferentes posições relativas ao ouvinte.

[0018] A FIG. 1 mostra um diagrama de blocos de alto nível do sintetizador de sinal binaural convencional 100, que converte um único sinal de fonte de áudio (ex., um sinal mono) em sinais de áudio esquerdo e direito de um sinal binaural, em que um sinal binaural é definido como sendo os dois sinais recebidos pelos tímpanos do ouvinte. Além do sinal fonte de áudio, o sintetizador 100 recebe um conjunto de sinalizações espaciais correspondentes à posição desejada da fonte de áudio relativa ao ouvinte. Nas implementações particulares, o conjunto de sinalizações espaciais compreende um valor de diferença de nível intercanais (ICLD) (que identifica a diferença no nível de áudio entre os sinais de áudio

esquerdo e direito da forma recebida pelos ouvidos esquerdo e direito, respectivamente) e um valor de diferença de tempo intercanais (ICTD) (que identifica a diferença de tempos da chegada dos sinais de áudio esquerdo e direito como recebidos pelos ouvidos esquerdo e direito, respectivamente). Além disso, ou como alternativa, algumas técnicas de síntese envolvem a modelagem de uma função de transferência dependente de direção para o som da fonte de sinais até os tímpanos, também denominada de função de transferência relativa à cabeça (HRTF). Ver, ex., J. Blauert, *The Psychophysics of Human Sound Localization*, MIT Press, 1983, cujos ensinamentos estão incorporados à presente como referência.

[0019] Usando o sintetizador binaural de sinais 100 da FIG. 1, o sinal de áudio mono gerado por uma única fonte sonora pode ser processado de maneira que, quando ouvido por meio dos fones de ouvido, a fonte sonora é localizada espacialmente pela aplicação de um conjunto adequado de sinalizações espaciais (ex., ICLD, ICTD, e/ou HRTF) para gerar o sinal de áudio para cada ouvido. Ver, ex., D. R. Begault, *3-D Sound for Virtual Reality and Multimedia*, Academic Press, Cambridge, Mass., 1994.

[0020] O sintetizador binaural de sinais 100 da FIG. 1 gera os tipos mais simples de **cenar auditivas**: aquelas que têm uma única fonte de áudio posicionada com relação ao ouvinte. **Cenar auditivas** mais complexas que compreendem duas ou mais fontes de áudio localizadas em diferentes posições relativas ao ouvinte podem ser geradas usando um sintetizador de **cena auditiva** que seja essencialmente implementado usando instâncias múltiplas de sintetizador binaural de sinais, em que cada instância de sintetizador binaural de sinais gera o sinal binaural que

corresponde a uma diferente fonte de áudio. Como cada fonte diferente de áudio tem uma localização diferente relativa ao ouvinte, é usado um diferente conjunto de sinalizações espaciais para gerar o sinal de áudio binaural para cada diferente fonte de áudio.

SUMÁRIO DA INVENÇÃO

[0021] De acordo com uma configuração, a presente invenção é um método, equipamento e meio de leitura por máquina para a codificação de canais de áudio. São gerados um ou mais códigos de sinalização para dois ou mais canais de áudio, caracterizado pelo fato de que pelo menos um código de sinalização é um código de sinalização combinado gerado pela combinação de dois ou mais códigos de sinalização estimados, e cada código de sinalização estimado é estimado a partir de um grupo de dois ou mais canais de áudio.

[0022] De acordo com uma outra configuração, a presente invenção é um equipamento para a codificação de C canais de áudio de entrada para gerar E canal(is) de áudio transmitido(s). O equipamento compreende um estimador de códigos e um downmixer. O estimador de códigos gera um ou mais códigos de sinalização para dois ou mais canais de áudio, caracterizado pelo fato de que pelo menos um código de sinalização é um código de sinalização combinado gerado pela combinação de dois ou mais códigos de sinalização estimados, e cada código de sinalização é estimado a partir de um grupo de dois ou mais canais de áudio. O downmixer faz o downmix dos C canais de entrada para gerar o(s) E canal(is) transmitido(s), em que $C > E \geq 1$, sendo o equipamento adaptado para transmitir as informações sobre os códigos de sinalização para permitir que o decodificador faça o processamento da síntese durante a decodificação do(s) E canal(is) transmitido(s).

[0023] De acordo com uma outra configuração, a presente

invenção é um fluxo de bits de áudio codificado gerado pela codificação de canais de áudio, caracterizado pelo fato de que um ou mais códigos de sinalização são gerados para dois ou mais canais de áudio, em que pelo menos um código de sinalização é um código de sinalização combinado gerado pela combinação de dois ou mais códigos de sinalização estimados, e cada código de sinalização estimado é estimado a partir de um grupo de dois ou mais canais de áudio. Os um ou mais códigos de sinalização e os E canal(is) de áudio transmitido(s) correspondente(s) a dois ou mais canais de áudio, em que $E \geq 1$, são codificado(s) no fluxo de bits de áudio codificado.

[0024] De acordo com uma outra configuração, a presente invenção é um fluxo de bits de áudio codificado compreendendo um ou mais códigos de sinalização e E canal(is) de áudio transmitido(s). Os um ou mais códigos de sinalização são gerados para dois ou mais canais de áudio, caracterizado pelo fato de que pelo menos um código de sinalização é um código de sinalização combinado gerado pela combinação de dois ou mais códigos de sinalização estimados e cada código de sinalização estimado é estimado a partir de um grupo de dois ou mais dos canais de áudio. O(s) E canal(is) de áudio transmitido(s) corresponde(m) aos dois ou mais canais de áudio.

[0025] De acordo com uma outra configuração, a presente invenção é um método, equipamento e meio de leitura por máquina para decodificação de E canal(is) de áudio transmitido(s) para gerar C canais de áudio playback, em que $C > E \geq 1$. São recebidos os códigos de sinalização correspondentes aos E canal(is) transmitido(s), caracterizado pelo fato de que pelo menos um código de sinalização é um código de sinalização combinado gerado pela combinação de dois ou mais códigos de sinalização estimados, e cada código de

sinalização estimado a partir de um grupo de dois ou mais canais de áudio correspondentes aos E canal(is) transmitido(s). Um ou mais dos E canal(is) transmitido(s) são upmixados para gerar um ou mais canais upmixados. Um ou mais dos C canais playback são sintetizados pela aplicação dos códigos de sinalização a um ou mais canais upmixados, caracterizado pelo fato de que dois ou mais códigos de sinalização derivados se derivam a partir do código de sinalização combinado, e cada código de sinalização derivado é aplicado para gerar dois ou mais canais sintetizados.

BREVE DESCRIÇÃO DOS DESENHOS

[0026] Outros aspectos, características e vantagens da presente invenção se tornarão completamente aparentes a partir da seguinte descrição detalhada, das reivindicações apenas e dos desenhos de acompanhamento em que números semelhantes de referência indicam elementos similares ou idênticos.

[0027] A FIG. 1 mostra um diagrama de blocos de alto nível do sintetizador binaural de sinais convencional;

[0028] A FIG. 2 é um diagrama de blocos de um sistema de processamento de áudio de codificação de sinalização binaural (*Binaural Cue Coding*) (BCC) genérico;

[0029] A FIG. 3 mostra um diagrama de blocos de um downmixer que pode ser usado para o downmixer da FIG. 2;

[0030] A FIG. 4 mostra um diagrama de blocos de um sintetizador BCC que pode ser usado para o decodificador da FIG. 2;

[0031] A FIG. 5 mostra um diagrama de blocos do estimador BCC da FIG. 2, de acordo com uma configuração da presente invenção;

[0032] A FIG. 6 ilustra a geração de dados ICTD e ICLD para o áudio de cinco canais;

[0033] A FIG. 7 ilustra a geração de dados ICC para o áudio de cinco canais;

[0034] A FIG. 8 mostra um diagrama de blocos de uma implementação do sintetizador BCC da FIG. 4 que pode ser usado em um decodificador BCC para gerar um sinal de áudio estéreo ou multicanais dado um único sinal de soma transmitido $s(n)$ mais as sinalizações espaciais;

[0035] A FIG. 9 ilustra como ICTD e ICLD variam dentro de uma sub-banda como uma função da frequência;

[0036] A FIG. 10 mostra um diagrama de blocos de um sintetizador BCC que pode ser usado para o decodificador da FIG. 2 em um esquema BCC 5-para-2; e

[0037] A FIG. 11 mostra um fluxograma do processamento de um sistema BCC, como o mostrado na FIG. 2, relacionado com uma configuração da presente invenção.

DESCRIÇÃO DETALHADA

[0038] Em *binaural cue coding* (BCC), um codificador codifica C canais de entrada de áudio para gerar E canais transmitidos de áudio, em que $C > E \geq 1$. Em particular, dois ou mais dos C canais de entrada são fornecidos no domínio de frequência, e um ou mais códigos de sinalização são gerados para cada um ou mais das diferentes bandas de frequência nos dois ou mais canais de entrada no domínio de frequência. Além disso, os C canais de entrada são downmixados para gerar os E canais transmitidos. Em algumas implementações de downmix, pelo menos um dos E canais transmitidos se baseia nos dois ou mais dos C canais de entrada, e pelo menos um dos E canais transmitidos se baseia em somente um dos C canais de entrada.

[0039] Em uma configuração, um codificador BCC tem dois ou mais bancos de filtros, um estimador de códigos e um downmixer. Os dois ou mais bancos de filtros convertem dois ou mais dos C canais de entrada de um domínio de tempo para um domínio de frequência. O estimador de códigos gera um ou mais códigos de sinalização para cada uma ou mais diferentes bandas de frequência nos dois ou mais canais de entrada convertidos. O downmixer faz o downmix dos C canais de entrada para gerar os E canais transmitidos, em que $C > E \geq 1$.

[0040] Na decodificação BCC, E canais transmitidos de áudio são decodificados para gerarem C canais de áudio playback. Em particular, para cada um ou mais diferentes bandas de frequência, um ou mais dos E canais transmitidos são upmixados em um domínio de frequência para gerar dois ou mais dos C canais playback no domínio de frequência, em que $C > E \geq 1$. Um ou mais códigos de sinalização são aplicados a cada um ou mais diferentes bandas de frequência nos dois ou mais canais playback no domínio de frequência para gerar dois ou mais canais modificados, e os dois ou mais canais modificados são convertidos do domínio de frequência para o domínio de tempo. Em algumas implementações de upmixing, pelo menos um dos C canais playback se baseia em pelo menos um dos E canais transmitidos e pelo menos um código de sinalização, e pelo menos um dos C canais playback se baseia em somente um dos E canais transmitidos e independente de quaisquer códigos de sinalização.

[0041] Em uma configuração, um decodificador BCC tem um upmixer, um sintetizador e um ou mais bancos de filtros inversos. Para cada uma ou mais diferentes bandas de frequência, o upmixer faz o upmix de um ou mais dos E canais transmitidos no domínio de frequência para gerar dois ou mais dos C canais playback no domínio

de frequência, em que $C \geq E \geq 1$. O sintetizador aplica um ou mais códigos de sinalização a cada uma ou mais diferentes bandas de frequência nos dois ou mais canais playback no domínio de frequência para gerar dois ou mais canais modificados. Um ou mais bancos de filtros inversos convertem os dois ou mais canais modificados do domínio de frequência em um domínio de tempo.

[0042] Dependendo da implementação em particular, um dado canal playback pode se basear em um único canal transmitido, ao invés de em uma combinação de dois ou mais canais transmitidos. Por exemplo, quando existir somente um canal transmitido, cada um dos C canais playback se baseia naquele canal transmitido. Nessas situações, o upmixing corresponde a copiar o canal transmitido correspondente. Assim, nas aplicações em que existe somente um canal transmitido, o upmixer pode ser implementado usando um replicador que copia o canal transmitido de cada canal playback.

[0043] Os codificadores e/ou decodificadores BCC podem ser incorporados a vários sistemas ou aplicações incluindo, por exemplo, gravadores/reprodutores de vídeo digital, gravadores/reprodutores de áudio digital, computadores, transmissores / receptores por satélite, transmissores / receptores a cabo, transmissores / receptores de difusão terrestre, sistemas de *home entertainment* e sistemas de *movie theater*.

Processamento BCC Genérico

[0044] A FIG. 2 é um diagrama de blocos de um sistema de processamento de áudio *binaural cue coding* (BCC) genérico 200 que compreende um codificador 202 e um decodificador 204. O codificador 202 inclui um downmixer 206 e um estimador BCC 208.

[0045] O downmixer 206 converte C canais de entrada de

áudio $x_i(n)$ em E canais transmitidos de áudio $y_i(n)$, em que $C > E \geq 1$. Nessa especificação, os sinais expressos usando a variável n são sinais no domínio de tempo, enquanto os sinais expressos usando a variável k são sinais no domínio de frequência. Dependendo da implementação em particular, o downmixing pode ser implementado tanto no domínio de tempo como no domínio de frequência. O estimador BCC 208 gera códigos BCC dos C canais de entrada de áudio e transmite aqueles códigos BCC tanto como informações auxiliares em banda como fora de banda relativas aos E canais transmitidos de áudio. Os códigos típicos BCC incluem um ou mais dos dados de diferenças de tempo intercanais (ICTD), diferenças de nível intercanais (ICLD) e correlação intercanais (ICC) estimados entre determinados pares de canais de entrada como uma função da frequência e do tempo. A implementação em particular determinará entre quais pares particulares de canais de entrada que os códigos BCC são estimados.

[0046] Os dados ICC correspondem à coerência de um sinal binaural, que se relaciona com a largura percebida da fonte de áudio. Quanto mais larga a fonte de áudio, mais baixa a coerência entre os canais esquerdo e direito do sinal binaural resultante. Por exemplo, a coerência do sinal binaural que corresponde à difusão de uma orquestra em um palco **auditivo** é tipicamente menor do que a coerência do sinal binaural que corresponde a um único violino tocando solo. Em geral, um sinal de áudio com menor coerência é normalmente percebido como mais difundido em um espaço **auditivo**. Assim, os dados ICC se relacionam tipicamente com a largura aparente da fonte e com o grau de envolvimento do ouvinte. Ver, por exemplo, J. Blauert, *The Psychophysics of Human Sound Localization*, MIT Press, 1983.

[0047] Dependendo da aplicação em particular, os E canais

de áudio transmitidos e os correspondentes códigos BCC podem ser transmitidos diretamente ao decodificador 204 ou armazenados em algum tipo adequado de dispositivo de armazenamento para acesso posterior pelo decodificador 204. Dependendo da situação, o termo "transmitindo" pode se referir tanto à transmissão direta a um decodificador ou armazenagem para a provisão subsequente a um decodificador. Em qualquer caso, o decodificador 204 recebe os canais de áudio transmitidos e as informações auxiliares e realiza o upmixing e a síntese BCC usando os códigos BCC para converter os E canais de áudio transmitidos em mais do que os E (tipicamente, mas não necessariamente, C) canais de áudio $\hat{x}_i(n)$ playback para playback de áudio. Dependendo da implementação em particular, o upmixing pode ser feito tanto no domínio de tempo como no domínio de frequência.

[0048] Além do processamento BCC mostrado na FIG. 2, um sistema genérico de processamento de áudio BCC pode incluir outros estágios de codificação e decodificação para comprimir mais os sinais de áudio no codificador, e então descomprimir os sinais de áudio no decodificador, respectivamente. Esses codecs de áudio podem se basear em técnicas de compressão/descompressão convencionais de áudio, como as baseadas em modulação em código de pulsos (PCM), PCM diferencial (DPCM) ou DPCM de adaptação (ADPCM).

[0049] Quando o downmixer 206 gera um sinal de soma simples (isto é, $E=1$), a codificação BCC pode representar sinais de áudio multicanais em um taxa de bits somente um pouco maior do que o necessário para representar um sinal de áudio mono. Isso ocorre porque os dados ICTD, ICLD e ICC estimados entre um par de canais contêm cerca de duas ordens de magnitude menos de informações do

que um formato de onda de áudio.

[0050] Não somente a baixa taxa de bits da codificação BCC, como também seu aspecto de retro-compatibilidade apresenta interesse. Um único sinal de soma transmitido corresponde a um downmix mono do sinal original estéreo ou multicanais. Para os receptores que não suportam reprodução sonora estéreo ou multicanais, ouvir o sinal de soma transmitido é um método válido de apresentação do material de áudio em um equipamento de reprodução mono de baixo desempenho. Portanto, a codificação BCC também pode ser usada para ampliar os serviços existentes que envolvem a difusão de material de áudio mono por áudio multicanais. Por exemplo, os sistemas existentes de radiodifusão de áudio mono podem ser ampliados para playback estéreo ou multicanais se as informações auxiliares BCC puderem ser integradas no canal existente de transmissão. Existem capacidades análogas ao ser feito o downmix de áudio multicanais para dois sinais de soma que correspondem ao áudio estéreo.

[0051] O BCC processa sinais de áudio com determinada resolução de tempo e frequência. A resolução de frequência usada é amplamente motivada pela resolução de frequência do sistema de audição humano. A psicoacústica sugere que a percepção espacial se baseia muito provavelmente em uma representação de banda crítica do sinal acústico de entrada. Essa resolução de frequência é considerada usando um banco de filtros inversível (por exemplo, baseado na transformada rápida de Fourier (FFT) ou um filtro espelhado em quadratura (QMF)) com sub-bandas com larguras de bandas iguais ou proporcionais à largura crítica de banda do sistema humano de audição.

Downmixing Genérico

[0052] Nas implementações preferidas, o(s) sinal(is) de soma transmitido(s) contém todas as componentes de sinais do sinal de áudio de entrada. O objetivo é que cada componente de sinal seja totalmente mantida. A simples soma dos canais de entrada de áudio geralmente resulta na amplificação ou na atenuação das componentes do sinal. Em outras palavras, a potência das componentes de sinal em uma soma "simples" é geralmente maior ou menor do que a soma da potência da componente de sinal correspondente de cada canal. Pode ser usada uma técnica de downmixing que equaliza o sinal de soma, de maneira que a potência das componentes de sinal no sinal de soma é aproximadamente a mesma que a potência correspondente em todos os canais de entrada.

[0053] A Fig. 3 mostra um diagrama de blocos de um downmixer 300 que pode ser usado para o downmixer 206 da FIG. 2 de acordo com certas implementações do sistema BCC 200. O downmixer 300 tem um banco de filtros (FB) 302 para cada canal de entrada $x_i(n)$, um bloco downmixing 304, um bloco opcional escalação/retardo 306 e um FB (IFB) inverso 308 para cada canal codificado $y_i(n)$.

[0054] Cada banco de filtros 302 converte cada frame (por exemplo, 20 msec) de um canal de entrada digital correspondente $x_i(n)$ no domínio de tempo em um conjunto de coeficientes de entrada $\tilde{x}_i(k)$ no domínio de frequência. O bloco de downmixing 304 faz o downmix de cada sub-banda de C coeficientes de entrada correspondentes em uma sub-banda correspondente de E coeficientes no domínio de frequência downmixados. A equação (1) representa o downmix da k^{a} sub-bandas de coeficientes de entrada $(\tilde{x}_1(k), \tilde{x}_2(k), \dots$

$\tilde{x}_c(k)$) para gerar a k^{a} sub-banda de coeficientes downmixados ($\hat{y}_1(k), \hat{y}_2(k), \dots, \hat{y}_E(k)$) como segue:

$$\begin{bmatrix} \hat{y}_1(k) \\ \hat{y}_2(k) \\ \vdots \\ \hat{y}_E(k) \end{bmatrix} = D_{CE} \begin{bmatrix} \tilde{x}_1(k) \\ \tilde{x}_2(k) \\ \vdots \\ \tilde{x}_C(k) \end{bmatrix},$$

(1)

[0055] onde D_{CE} é uma matriz de downmixing de valores reais C-por-E.

[0056] O bloco de escalação/retardo opcional 306 compreende um conjunto de multiplicadores 310, cada um dos quais multiplica um coeficiente downmixado correspondente $\hat{y}_i(k)$ por um fator de escalação $e_i(k)$ para gerar um coeficiente escalado correspondente $\tilde{y}_i(k)$. A motivação para a operação de escalação é equivalente à equalização generalizada para downmixing com fatores arbitrários de ponderação para cada canal. Se os canais de entrada forem independentes, então a potência $p_{\tilde{y}_i(k)}$ do sinal downmixado em cada sub-banda é dada pela Equação (2) como segue:

$$\begin{bmatrix} p_{\tilde{y}_1(k)} \\ p_{\tilde{y}_2(k)} \\ \vdots \\ p_{\tilde{y}_E(k)} \end{bmatrix} = \overline{D}_{CE} \begin{bmatrix} p_{\tilde{x}_1(k)} \\ p_{\tilde{x}_2(k)} \\ \vdots \\ p_{\tilde{x}_C(k)} \end{bmatrix},$$

(2)

[0057] onde \overline{D}_{CE} se deriva pela quadratura de cada elemento de matriz na matriz de downmixing C-por-E D_{CE} e $p_{\tilde{x}_i(k)}$ é a potência da sub-banda k do canal de entrada i .

[0058] Se as sub-bandas não forem independentes, então os

valores de potência $p_{\tilde{y}_i(k)}$ do sinal downmixado serão maiores ou menores do que os computados usando a Equação (2), devido às amplificações ou cancelamentos de sinais quando as componentes do sinal estiverem em fase ou fora de fase, respectivamente. Para evitar isso, a operação de downmixing da Equação (1) se aplica a sub-bandas seguidas pela operação de escalação dos multiplicadores 310. Os fatores de escalação $e_i(k)$ ($1 \leq i \leq E$) podem ser obtidos usando a Equação (3) como segue:

$$e_i(k) = \sqrt{\frac{p_{\tilde{y}_i(k)}}{p_{\hat{y}_i(k)}}}, \quad (3)$$

[0059] onde $p_{\tilde{y}_i(k)}$ é a potência de sub-banda computada pela Equação (2), e $p_{\hat{y}_i(k)}$ é a potência do sinal de sub-banda downmixado correspondente $\hat{y}_i(k)$.

[0060] Além de, ou ao invés de prover escalação opcional, o bloco de escalação/retardo 306 pode aplicar opcionalmente retardos aos sinais.

[0061] Cada banco de filtros inverso 308 converte um conjunto de coeficientes escalados correspondentes $\tilde{y}_i(k)$ do domínio de frequência para um frame de um canal transmitido digital correspondente $y_i(n)$.

[0062] Apesar de a FIG. 3 mostrar todos os C canais de entrada sendo convertidos para o domínio de frequência para subsequente downmixing, em implementações alternativas, um ou mais (mas não menos do que $C-1$) dos C canais de entrada podem desviar alguns ou todos os processamentos mostrados na FIG. 3 e ser transmitidos como um número equivalente de canais de áudio não

modificados. Dependendo da implementação em particular, esses canais de áudio não modificados podem ou não ser usados pelo estimador BCC 208 da FIG. 2 na geração dos códigos BCC transmitidos.

[0063] Em uma implementação do downmixer 300 que gera um sinal de soma simples $y(n)$, $E=1$ e os sinais $\tilde{x}_c(k)$ de cada sub-banda de cada canal de entrada c são adicionados e então multiplicados por um fator $e(k)$, de acordo com a Equação (4) como segue:

$$\tilde{y}(k) = e(k) \sum_{c=1}^C \tilde{x}_c(k). \quad (4)$$

[0064] o fator $e(k)$ é dado pela Equação (5) como segue:

$$e(k) = \sqrt{\frac{\sum_{c=1}^C p_{\tilde{x}_c}(k)}{p_{\tilde{x}}(k)}}, \quad (5)$$

[0065] onde $p_{\tilde{x}_c}(k)$ é uma breve estimativa da potência de $\tilde{x}_c(k)$ no índice de tempo k , e $p_{\tilde{x}}(k)$ é uma breve estimativa da potência de $\sum_{c=1}^C \tilde{x}_c(k)$. As sub-bandas equalizadas são transformadas novamente para o domínio de tempo resultando no sinal de soma $y(n)$ que é transmitido ao decodificador BCC.

Síntese BCC Genérica

[0066] A FIG. 4 mostra um diagrama de blocos de um sintetizador BCC 400 que pode ser usado para o decodificador 204 da FIG. 2 de acordo com certas implementações do sistema BCC 200. O sintetizador BCC 400 tem um banco de filtros 402 para cada canal transmitido $y_i(n)$, um bloco de upmixing 404, de retardos 406, de multiplicadores 408, de blocos de correlação 410, e um banco de filtros inverso 412 para cada canal playback $\hat{x}_i(n)$.

[0067] Cada banco de filtros 402 converte cada frame de um canal digital transmitido correspondente $y_i(n)$ do domínio de tempo em um conjunto de coeficientes de entrada $\tilde{y}_i(k)$ no domínio de frequência. O bloco de upmixing 404 faz o upmix de cada sub-banda dos E coeficientes de canal transmitidos correspondentes em uma sub-banda correspondente de C coeficientes upmixados no domínio de frequência. A equação (4) representa o upmixing da k° sub-banda de coeficientes de canais transmitidos $(\tilde{y}_1(k), \tilde{y}_2(k), \dots, \tilde{y}_E(k))$ para gerar a k° sub-banda de coeficientes upmixados $(\tilde{s}_1(k), \tilde{s}_2(k), \dots, \tilde{s}_C(k))$ como segue:

$$\begin{bmatrix} \tilde{s}_1(k) \\ \tilde{s}_2(k) \\ \vdots \\ \tilde{s}_C(k) \end{bmatrix} = U_{EC} \begin{bmatrix} \tilde{y}_1(k) \\ \tilde{y}_2(k) \\ \vdots \\ \tilde{y}_E(k) \end{bmatrix}, \quad (6)$$

[0068] onde U_{EC} é uma matriz de upmixing de valores reais E -por- C . A realização do upmixing no domínio de frequência permite que o upmixing seja aplicado individualmente em cada diferente sub-banda.

[0069] Cada retardo 406 aplica um valor de retardo $d_i(k)$ com base no código BCC correspondente para dados ICTD para garantir que os valores ICTD desejados apareçam entre certos pares de canais playback. Cada multiplicador 408 aplica um fator de escalação $a_i(k)$ com base no código BCC correspondente para dados ICLD para garantir que os valores ICLD desejados apareçam entre certos pares de canais playback. O bloco de correlação 410 realiza uma operação de decorrelação A com base nos correspondentes códigos BCC para dados ICC para garantir que os valores ICC desejados apareçam entre certos pares de canais playback. Outras descrições das operações de bloco

de correlação 410 podem ser encontradas no Pedido de Patente Norte-Americana Nº 10/155.437, depositado em 24 de maio de 2002 como Baumgarte 2-10.

[0070] A síntese dos valores ICLD pode ser menos problemática do que a síntese dos valores ICTD e ICC, já que a síntese ICLD envolve somente a escalação de sinais de sub-banda. Como as sinalizações ICLD são as sinalizações direcionais mais comumente usadas, é normalmente mais importante que os valores ICLD se aproximem daqueles do sinal de áudio original. Assim, os dados ICLD podem ser estimados entre todos os pares de canais. Os fatores de escalação $a_i(k)$ ($1 \leq i \leq C$) para cada sub-banda são escolhidos de preferência de maneira que a potência de sub-banda de cada canal playback se aproxime da potência correspondente do canal de entrada de áudio original.

[0071] Um objetivo pode ser aplicar relativamente poucas modificações de sinal para a sintetização dos valores ICTD e ICC. Assim, os dados BCC podem não incluir valores ICTD e ICC para todos os pares de canais. Nesse caso, o sintetizador BCC 400 sintetizaria os valores ICTD e ICC somente entre determinados pares de canais.

[0072] Cada banco de filtro inverso 412 converte um conjunto de coeficientes sintetizados correspondentes $\tilde{x}_i(k)$ do domínio de frequência para um frame de um correspondente canal digital de playback $\hat{x}_i(n)$.

[0073] Apesar de a FIG. 4 mostrar todos os E canais transmitidos sendo convertidos para o domínio de frequência para subsequente upmixing e processamento BCC, nas implementações alternativas, um ou mais (mas não todos) dos E canais transmitidos

pode(m) desviar de alguns ou todos os processamentos mostrados na Fig. 4. Por exemplo, um ou mais dos canais transmitidos podem ser canais não modificados que não estejam submetidos a qualquer upmixing. Além de serem um ou mais dos C canais playback, esses canais não modificados, por sua vez podem, mas não precisam ser usados como canais de referência para os quais o processamento BCC é aplicado para sintetizar um ou mais de outros canais playback. Em qualquer caso, esses canais não modificados podem estar sujeitos a retardos para compensar o tempo de processamento envolvido no upmixing e/ou no processamento BCC usado para gerar o restante dos canais playback.

[0074] Notar que, apesar de a FIG. 4 mostrar C canais playback sendo sintetizados a partir de E canais transmitidos, em que C é também o número de canais de entrada originais, a síntese BCC não se limita a esse número de canais playback. Em geral, o número de canais playback pode ser qualquer número de canais, incluindo números maiores ou menores que C e possivelmente mesmo situações em que o número de canais playback é igual ou menor que o número de canais transmitidos.

"Diferenças Perceptivelmente Relevantes" Entre
Canais de Áudio

[0075] Supondo um sinal de soma único, BCC sintetiza um sinal de áudio estéreo ou multicanais de maneira que ICTD, ICLD e ICC se aproximam das sinalizações correspondentes do sinal de áudio original. A seguir, será discutido o papel de ICTD, ICLD e ICC com relação aos atributos de imagem espacial **auditiva**.

[0076] O conhecimento sobre audição espacial implica que, para um evento **auditivo**, ICTD e ICLD estão relacionados com a direção

percebida. Ao se considerar as respostas de impulsos binaurais do ambiente (BRIRs) de uma fonte, existe uma relação entre a largura do evento **auditivo** e o envolvimento do ouvinte e os dados ICC estimados para as partes anteriores e posteriores das BRIRs. Entretanto, a relação entre ICC e essas propriedades de sinais gerais (e não somente as BRIRs) não é direta.

[0077] Sinais de áudio estéreo e multicanais normalmente contêm uma mistura complexa de sinais fonte concorrentemente ativos superpostos pelas componentes do sinal refletido, resultando da gravação em espaços fechados ou adicionados pelo engenheiro de gravação para a criação artificial de uma impressão espacial. Diferentes sinais fonte e suas reflexões ocupam diferentes regiões no plano tempo-frequência. Isto é refletido pelas ICTD, ICLD e ICC, que variam como uma função do tempo e da frequência. Nesse caso, a relação entre as ICTD, ICLD e ICC instantâneas e as direções do evento **auditivo** e a impressão espacial não é óbvia. A estratégia de determinadas configurações de BCC é sintetizar cegamente essas sinalizações, de maneira que se aproximem das sinalizações correspondentes do sinal de áudio original.

[0078] São usados bancos de filtros com sub-bandas de larguras de bandas iguais a duas vezes a largura de banda retangular equivalente (ERB). A audição informal revela que a qualidade de áudio da BCC não melhora de maneira notável ao ser escolhida uma maior resolução de frequência. Pode ser desejada uma menor resolução de frequência, já que isso resulta em menos valores ICTD, ICLD e ICC que devem ser transmitidos ao decodificador e assim em menor taxa de bits.

[0079] Com relação à resolução de tempo, ICTD, ICLD e ICC

são consideradas tipicamente em intervalos regulares de tempo. É obtido alto desempenho quando ICTD, ICLD e ICC são considerados aproximadamente a cada 4 a 16 ms. Notar que, a menos que as sinalizações sejam consideradas em intervalos de tempo muito pequenos, o efeito de precedência não é considerado diretamente. Supondo um par *lead-lag* clássico de estímulos sonoros, se o *lead* e o *lag* caírem em um intervalo de tempo em que seja sintetizado somente um conjunto de sinalizações, então a dominância de localização do *lead* não será considerada. Apesar disso, BCC obtém qualidade de áudio refletida em uma classificação MUSHRA média de cerca de 87 (isto é, qualidade "excelente" de áudio) na média e até perto de 100 para determinados sinais de áudio.

[0080] A diferença perceptivelmente pequena geralmente obtida entre o sinal de referência e o sinal sintetizado implica que as sinalizações relacionadas a uma ampla faixa de atributos de imagem espacial **auditivo** são consideradas implicitamente pela sintetização de ICTD, ICLD e ICC em intervalos regulares de tempo. A seguir, são feitas algumas argumentações sobre como ICTD, ICLD e ICC que podem se relacionar com uma faixa de atributos de imagem espacial **auditivo**.

Estimativa de Sinalizações Espaciais

[0081] A seguir é descrito como são estimadas ICTD, ICLD e ICC. A taxa de bits para a transmissão dessas (quantizadas e codificadas) sinalizações espaciais pode ser de somente alguns kb/s e portanto, com BCC, é possível transmitir sinais de áudio estéreo e multicanais em taxas de bits próximas às necessárias para um único canal de áudio.

[0082] A FIG. 5 mostra um diagrama de blocos de um

estimador BCC 208 da FIG. 2, de acordo com uma configuração da presente invenção. O estimador BCC 208 compreende bancos de filtros (FB) 502, que podem ser os mesmos bancos de filtros 302 da FIG. 3, e o bloco de estimativa 504, que gera sinalizações espaciais ICTD, ICLD e ICC para cada diferente sub-banda de frequência gerada pelos bancos de filtros 502.

Estimativa de ICTD, ICLD e ICC para Sinais Estéreo

[0083] As seguintes medidas são usadas para ICTD, ICLD e ICC para os correspondentes sinais de sub-banda $\tilde{x}_1(k)$ e $\tilde{x}_2(k)$ de dois canais de áudio (por exemplo, estéreo):

[0084] ICTD [amostras]:

$$\tau_{12}(k) = \arg \max_d \{\Phi_{12}(d, k)\}, \quad (7)$$

[0085] com uma breve estimativa da função normalizada de correlação cruzada dada pela Equação (8) como segue:

$$\Phi_{12}(d, k) = \frac{p_{\tilde{x}_1\tilde{x}_2}(d, k)}{\sqrt{p_{\tilde{x}_1}(k-d_1)p_{\tilde{x}_2}(k-d_2)}}, \quad (8)$$

onde

$$\begin{aligned} d_1 &= \max\{-d, 0\} \\ d_2 &= \max\{d, 0\}, \end{aligned} \quad (9)$$

e $p_{\tilde{x}_1\tilde{x}_2}(d, k)$ é uma breve estimativa da média de $\tilde{x}_1(k-d_1)\tilde{x}_2(k-d_2)$.

o ICLD [dB]:

$$\Delta L_{12}(k) = 10 \log_{10} \left(\frac{p_{\tilde{x}_2}(k)}{p_{\tilde{x}_1}(k)} \right). \quad (10)$$

o ICC:

$$c_{12}(k) = \max_d |\Phi_{12}(d, k)|. \quad (11)$$

[0086] Notar que é considerado o valor absoluto da correlação cruzada normalizada e $c_{12}(k)$ tem uma faixa de $[0,1]$.

Estimativa de ICTD, ICLD e ICC para Sinais de Áudio Multicanais

[0087] Quando existirem mais do que dois canais de entrada, é normalmente suficiente definir ICTD e ICLD entre um canal de referência (por exemplo, canal número 1) e os demais canais, como ilustrado na Fig. 6 para o caso de $C=5$ canais, em que $\tau_{1c}(k)$ e $\Delta L_{1c}(k)$ denotam ICTD e ICLD, respectivamente, entre o canal de referência 1 e o canal c .

[0088] De forma oposta à ICTD e ICLD, ICC normalmente tem mais graus de liberdade. Os ICC como definidos, podem ter valores diferentes entre todos os possíveis pares de canais de entrada. Para C canais, existem $C(C-1)/2$ possíveis pares de canais; por exemplo, para 5 canais, existem 10 pares de canais como ilustrado na Fig. 7(a). Entretanto, tal esquema exige que, para cada sub-banda em cada índice de tempo, os valores ICC $C(C-1)/2$ sejam estimados e transmitidos, resultando em alta complexidade de computação e em alta taxa de bits.

[0089] De maneira alternativa, para cada sub-banda, ICTD e ICLD determinam a direção na qual o evento **auditivo** da componente de sinal correspondente é reproduzido na sub-banda. Um único parâmetro ICC por sub-banda pode então ser usado para descrever a coerência geral entre todos os canais de áudio. Podem ser obtidos bons resultados pela estimativa e transmissão de sinalizações ICC somente entre os dois canais com maior energia em cada sub-banda em

cada índice de tempo. Isto está ilustrado na Fig. 7(b), em que para os instantes de tempo $k-1$ e k os pares de canais (3, 4) e (1, 2) são os mais fortes, respectivamente. Pode ser usada uma regra heurística para determinar ICC entre os demais pares de canais.

Síntese de Sinalizações Espaciais

[0090] A Fig. 8 mostra um diagrama de blocos de uma implementação do sintetizador BCC 400 da FIG. 4 que pode ser usado em um decodificador BCC para gerar um sinal de áudio estéreo ou multicanais dado um único sinal de soma transmitido $s(n)$ mais as sinalizações espaciais. O sinal de soma $s(n)$ é decomposto em sub-bandas, em que $\tilde{s}(k)$ denota uma dessas sub-bandas. Para a geração das correspondentes sub-bandas de cada um dos canais de saída, são aplicados retardos d_c , fatores de escala a_c , e filtros h_c à sub-banda correspondente do sinal de soma. (Para a simplicidade de notação, o índice de tempo k é ignorado nos retardos, nos fatores de escala e nos filtros.) As ICTD são sintetizadas por retardos impostos, as ICLD por escalação e ICC pela aplicação de filtros de decorrelação. O processamento mostrado na Fig. 8 se aplica de forma independente a cada sub-banda.

Síntese ICTD

[0091] Os retardos d_c são determinados a partir das ICTDs $\tau_{1c}(k)$, de acordo com a Equação (12) como a seguir:

$$\begin{aligned} d_c &= -\frac{1}{2}(\max_{2 \leq l \leq C} \tau_{1l}(k) + \min_{2 \leq l \leq C} \tau_{1l}(k)), \text{ for } c = 1 \\ d_c &= \tau_{1l}(k) + d_1, \text{ for } 2 \leq c \leq C. \end{aligned} \quad (12)$$

[0092] O retardo para o canal de referência, d_1 , é computado de maneira que a magnitude máxima dos retardos d_c seja minimizada. Quanto menos os sinais de sub-banda forem modificados,

menor é o perigo de ocorrência de problemas. Se a taxa de amostragem de sub-banda não proporcionar suficiente tempo-resolução para síntese ICTD, os retardos poderão ser impostos com maior precisão usando os filtros passa-tudo adequados.

Síntese ICLD

[0093] Para que os sinais de sub-banda de saída tenham os desejados ICLDs, o $\Delta L_{12}(k)$ entre o canal c e o canal de referência 1 e os fatores de ganho a_c devem satisfazer a Equação (13) como segue:

$$\frac{a_c}{a_1} = 10^{\frac{\Delta L_{1c}(k)}{20}}. \quad (13)$$

[0094] Além disso, as sub-bandas de saída são preferivelmente normalizadas de maneira que a soma da potência de todos os canais de saída seja igual à potência do sinal de soma de entrada. Como a potência do sinal original total de cada sub-banda é preservada no sinal de soma, essa normalização resulta na potência de sub-banda absoluta de cada canal de saída aproximando-se da potência correspondente do sinal de áudio de entrada do codificador original. Dadas essas limitações, os fatores de escala a_c são dados pela Equação (14) como segue:

$$\begin{aligned} a_c &= 1/\sqrt{1 + \sum_{i=2}^C 10^{\Delta L_{1i}/10}}, \text{ para } c = 1 \\ a_c &= 10^{\Delta L_{1c}/20} a_1, \text{ ou } (c \neq 1). \end{aligned} \quad (14)$$

Síntese ICC

[0095] Em certas configurações, o objetivo da síntese ICC é reduzir a correlação entre as sub-bandas após os retardos e tendo sido aplicada a escalação, sem afetar ICTD e ICLD. Isto pode ser conseguido projetando os filtros h_c na Fig. 8 de maneira que ICTD e ICLD variem efetivamente como uma função da frequência, de maneira

que a variação média seja zero em cada sub-banda (banda crítica **auditiva**).

[0096] A Fig. 9 ilustra como ICTD e ICLD variam dentro de uma sub-banda como uma função da frequência. A amplitude da variação de ICTD e ICLD determina o grau de decorrelação e é controlada como uma função de ICC. Notar que ICTD varia suavemente (como na Fig. 9(a)), enquanto ICLD varia randomicamente (como na Fig. 9(b)). Pode-se variar ICLD tão suavemente quanto ICTD, mas isto resultaria em uma maior coloração dos sinais de áudio resultantes.

[0097] Outro método para sintetizar ICC, particularmente adequado para a síntese ICC multicanais, é descrita em maiores detalhes em C. Faller, "Parametric multi-channel audio coding: Synthesis of coherence cues," IEEE Trans. on Speech and Audio Proc., 2003, cujos ensinamentos estão incorporados à presente como referência. Como uma função do tempo e da frequência, quantidades específicas de reverberação tardia artificial são adicionadas a cada um dos canais de saída para a obtenção do ICC desejado. Além disso, pode ser aplicada uma modificação espectral, de maneira que o envelope espectral do sinal resultante se aproxime do envelope espectral do sinal de áudio original.

[0098] Outras técnicas de síntese ICC relativas e não relativas a sinais estéreo (ou pares de canais de áudio) foram apresentadas em E. Schuijers, W. Oomen, B. den Brinker, and J. Breebaart, "Advances in parametric coding for high-quality audio," in Preprint 114th Conv. Aud. Eng. Soc., March 2003, and J. Engdegard, H. Purnhagen, J. Roden, and L. Liljeryd, "Synthetic ambience in parametric stereo coding," in Preprint 117th Conv. Aud. Eng. Soc., May 2004, cujos ensinamentos estão incorporados à presente como

referência.

BCC C-para-E

[0099] Como descrito anteriormente, o BCC pode ser implementado com mais do que um canal de transmissão. Foi descrita uma variação de BCC que representa *C* canais de áudio não como um único canal (transmitido), mas como canais *E*, indicados como BCC *C-para-E*. Existem (pelo menos) duas motivações para BCC *C-para-E*:

[00100] BCC com um canal de transmissão provê um caminho retro-compatível para a atualização dos sistemas mono existentes para playback de áudio estéreo ou multicanais. Os sistemas atualizados transmitem o sinal soma downmixado BCC pela infraestrutura mono existente, enquanto também transmite as informações auxiliares BCC. BCC *C-para-E* se aplica à codificação retro-compatível dos canais *E* de áudio do canal *C*.

[00101] BCC *C-para-E* introduz escalabilidade em termos de diferentes graus de redução do número de canais transmitidos. Espera-se que quanto mais canais de áudio sejam transmitidos, melhor será a qualidade de áudio. Os detalhes do processamento de sinal para BCC *C-para-E*, de maneira a definir as sinalizações ICTD, ICLD e ICC estão descritos no pedido norte-americano número de série 10/762.100, depositado em 20 de janeiro de 2004 (Faller 13-1).

Informações Auxiliares Compactas

[00102] Como acima descrito, em um esquema BCC típico, o codificador transmite ao decodificador códigos ICTD, ICLD e/ou ICC estimados entre diferentes pares ou grupos de canais de áudio. Essas informações auxiliares são transmitidas em adição ao(s) sinal(is) (ex., mono ou estéreo) downmix de maneira a obter um sinal de áudio multicanais após a decodificação BCC. Assim, é desejável minimizar

a quantidade de informações auxiliares, enquanto não se degrada a qualidade subjetiva do som decodificado.

[00103] Como os valores ICLD e ICTD se relacionam tipicamente com um canal de referência, os valores ICLD e ICTD C-1 são suficientes para descrever as características dos C canais codificados. Por outro lado, os ICCs são definidos entre pares arbitrários de canais. Assim, para C canais codificados, existem $C(C-1)/2$ possíveis pares ICC. Para 5 canais codificados, isto corresponde a 10 pares ICC. Na prática, para limitar a quantidade de informações ICC transmitidas, somente são transmitidas as informações ICC de determinados pares.

[00104] A FIG. 10 mostra um diagrama de blocos de um sintetizador BCC 1000 que pode ser usado para o decodificador 204 da FIG. 2 em um esquema BCC 5-para-2. Como mostrado na FIG. 10, o sintetizador BCC 1000 recebe dois sinais de entrada $y_1(n)$ e $y_2(n)$ e informações auxiliares BCC (não mostradas) gerando cinco sinais sintetizados de saída $\hat{x}_1(n), \dots, \hat{x}_5(n)$, em que o primeiro, o segundo, o terceiro, o quarto e o quinto sinais de saída correspondem aos sinais surround esquerdo, direito, central, esquerdo traseiro e direito traseiro, respectivamente, mostrados nas FIGS. 6 e 7.

[00105] Os parâmetros de retardo, escalação e decorrelação derivados das informações auxiliares transmitidas ICTD, ICLD e ICC são aplicados nos elementos 1004, 1006 e 1008, respectivamente, para sintetizar os cinco sinais de saída $\hat{x}_i(n)$ dos cinco sinais " upmixados " $\tilde{s}_i(k)$ gerados pelo elemento de upmixing 1002. Como mostrado na FIG. 10, a decorrelação somente é feita entre os canais esquerdo e traseiro esquerdo (isto é, canais 1 e 4) e entre os canais

direito e traseiro direito (isto é, canais 2 e 5). Assim, não mais do que dois conjuntos de dados ICC devem ser transmitidos ao sintetizador BCC 1000, no qual esses dois conjuntos caracterizam os valores ICC entre os dois pares de canais de cada sub-banda. Apesar de esta já ser uma redução considerável da quantidade de informações auxiliares ICC, é desejável mais uma redução.

[00106] De acordo com uma configuração da presente invenção, no contexto do esquema BCC 5-para-2 da FIG. 10, para cada sub-banda, o codificador BCC correspondente combina o valor ICC estimado para o par de canais "esquerdo/traseiro esquerdo" com o valor ICC estimado para o par de canais "direito/traseiro direito" para gerar um único valor ICC combinado que efetivamente indica uma quantidade global de decorrelação frontal/traseira e que é transmitida para o decodificador BCC como informações auxiliares ICC. Experimentos informais indicaram que esta simplificação não resulta virtualmente em nenhuma perda da qualidade do áudio, reduzindo pela metade as informações ICC transmitidas.

[00107] Em geral, as configurações da presente invenção são direcionadas para esquemas BCC, em que dois ou mais diferentes ICCs estimados entre diferentes pares de canais, ou grupos de canais, são combinados para transmissão, como indicado pela Equação (15) seguinte:

$$ICC_{transmitted} = f(ICC_1, ICC_2, \dots, ICC_N), \quad (15)$$

[00108] onde f é uma função que combina N diferentes ICCs.

[00109] Para obter uma medida ICC combinada que seja representativa da imagem espacial, pode ser vantajoso usar uma média ponderada para a função f que considere a importância dos canais individuais, na qual a importância do canal pode se basear nas

potências dos canais, como representado pela Equação (16) a seguir:

$$ICC_{transmitted} = \frac{\sum_{i=1}^N p_i ICC_i}{\sum_{i=1}^N p_i}, \quad (16)$$

[00110] em que p_i é a potência do par de canais correspondente na sub-banda. Nesse caso, os ICCs estimados a partir dos pares de canais mais potentes são ponderados mais do que estimados por ICCs a partir dos pares de canais mais fracos. A potência combinada p_i de um par de canais pode ser computada como a soma das potências dos canais individuais em cada sub-banda.

[00111] No decodificador, os $ICC_{transmitted}$, ICCs dados podem ser derivados para cada par de canais. Em uma possível implementação, o decodificador simplesmente usa $ICC_{transmitted}$ como o código ICC derivado para cada par de canais. Por exemplo, no contexto do esquema BCC 5-para-2 da FIG. 110, $ICC_{transmitted}$ pode ser usado diretamente para a decorrelação tanto do par de canais esquerdo/traseiro esquerdo como do par de canais direito/traseiro direito.

[00112] Em uma outra possível implementação, se o decodificador estimar as potências do par de canais a partir dos sinais sintetizados, então a ponderação da Equação (16) pode ser estimada e o processo decodificador pode usar opcionalmente estas informações e outros argumentos estatísticos perceptivos e de sinais para a geração de uma regra para a derivação de dois códigos ICC individuais e perceptivamente otimizados.

[00113] Apesar de a combinação dos valores ICC ter sido descrita no contexto de um determinado esquema BCC 5-para-2, a presente invenção pode ser implementada no contexto de qualquer esquema BCC C-para-E, incluindo aqueles em que $E=1$.

[00114] A FIG. 11 mostra um fluxograma do processamento de

um sistema BCC, como o mostrado na FIG. 2, relacionado com uma configuração da presente invenção. A FIG. 11 mostra somente as etapas associadas com o processamento relativo ao ICC.

[00115] Em particular, um codificador BCC estima os valores ICC entre dois ou mais grupos de canais (etapa 1102), combina dois ou mais daqueles valores ICC estimados para gerar um ou mais valores ICC combinados (etapa 1104), e transmite os valores ICC combinados (possivelmente em conjunto com um ou mais valores ICC "descombinados") como informações auxiliares BCC para um decodificador BCC (etapa 1106). O decodificador BCC deriva dois ou mais valores ICC dos valores ICC combinados recebidos (etapa 1108) e decorrelaciona grupos de canais usando os valores ICC derivados (e possivelmente um ou mais valores ICC não combinados recebidos) (etapa 1110).

Outras Configurações Alternativas

[00116] A presente invenção foi descrita no contexto do esquema BCC 5-para-2 da FIG. 10. Naquele exemplo, um codificador BCC (1) estima dois códigos ICC para dois pares de canais que consistem de quatro canais diferentes (isto é, esquerdo/traseiro esquerdo e direito/traseiro direito) e (2) media esses dois códigos ICC para gerar um código ICC combinado, que é transmitido para um decodificador BCC. O decodificador BCC (1) deriva dois códigos ICC do código ICC combinado transmitido (notar que o código ICC combinado pode ser simplesmente usado para ambos os códigos ICC derivados) e (2) aplica cada um dos dois códigos ICC derivados a um diferente par de canais sintetizados para gerar quatro canais decorrelacionados (isto é, canais sintetizados esquerdo, traseiro esquerdo, direito e traseiro direito).

[00117] A presente invenção também pode ser implementada em outros contextos. Por exemplo, um codificador BCC poderia estimar dois códigos ICC a partir de três canais de entrada A, B e C, em que um código ICC estimado corresponde aos canais A e B, e o outro código ICC estimado corresponde aos canais A e C. Nesse caso, o codificador poderia ser dito como estimativo de dois códigos ICC a partir de dois pares de canais de entrada, em que os dois pares de canais de entrada compartilham um canal comum (isto é, o canal de entrada A). O codificador poderia então gerar e transmitir um único código ICC combinado baseado nos dois códigos ICC estimados. Um decodificador BCC poderia então derivar dois códigos ICC a partir do código ICC combinado transmitido e aplicar esses dois códigos ICC derivados para sintetizar três canais decorrelacionados (isto é, canais sintetizados A, B e C). Nesse caso, cada código ICC derivado pode ser dito como aplicado para gerar um par de canais decorrelacionados, em que os dois pares de canais decorrelacionados compartilham um canal comum (isto é, o canal sintetizado A).

[00118] Apesar de a presente invenção ter sido descrita no contexto de esquemas de codificação BCC que empregam códigos ICC combinados, a presente invenção também pode ser implementada no contexto de esquemas de codificação BCC que empregam códigos de sinalização BCC combinados gerados pela combinação de dois ou mais códigos de sinalização BCC além dos códigos ICC, como os códigos ICTD e/ou os códigos ICLD, ao invés de, ou além do emprego dos códigos ICC combinados.

[00119] Apesar de a presente invenção ter sido descrita no contexto dos esquemas de codificação BCC envolvendo os códigos ICTD, ICLD e ICC, a presente invenção também pode ser implementada no

contexto de outros esquemas de codificação BCC envolvendo somente um ou dois desses três tipos de códigos (ex., ICLD e ICC, mas não ICTD) e/ou um ou mais tipos adicionais de códigos.

[00120] No esquema BCC 5-para-2 representado na FIG. 10, os dois canais transmitidos $y_1(n)$ e $y_2(n)$ são tipicamente gerados pela aplicação de um determinado esquema de estágio único de downmixing aos cinco canais mostrados nas FIGS. 6 e 7, em que o canal y_1 é gerado como uma soma ponderada dos canais 1, 3 e 4, e o canal y_2 é gerado como uma soma ponderada dos canais 2, 3 e 5, em que, por exemplo, em cada soma ponderada, o fator ponderado do canal 3 é a metade do fator ponderado usado para cada um dos dois outros canais. Nesse esquema BCC de estágio único, os códigos de sinalização BCC estimados correspondem a diferentes pares dos cinco canais originais de entrada. Por exemplo, um conjunto de códigos ICC estimados se baseia nos canais 1 e 4 e outro conjunto de códigos ICC estimados se baseia nos canais 2 e 5.

[00121] Em um esquema BCC alternativo de estágios múltiplos, os canais são downmixados sequencialmente, com os códigos de sinalização BCC correspondendo potencialmente a diferentes grupos de canais em cada estágio da seqüência de downmixing. Por exemplo, para os cinco canais das FIGS. 6 e 7 de um codificador BCC, os canais esquerdo e traseiro esquerdo originais podem ser downmixados para formarem um primeiro canal esquerdo downmixado com um primeiro conjunto de códigos de sinalização BCC gerados correspondendo àqueles dois canais originais. De maneira similar, os canais direito e traseiro direito originais poderiam ser downmixados para formarem um primeiro canal direito downmixado com um segundo conjunto de códigos de sinalização BCC gerados correspondendo àqueles dois

canais originais. Em um segundo estágio de downmixing, o primeiro canal esquerdo downmixado poderia ser downmixado com o canal central original para formarem um segundo canal esquerdo downmixado com um terceiro conjunto de códigos de sinalização BCC gerados correspondendo ao primeiro canal esquerdo downmixado e o canal central original. De maneira similar, o primeiro canal direito downmixado poderia ser downmixado com o canal central original para formarem um segundo canal direito downmixado com um quarto conjunto de códigos de sinalização BCC gerados correspondendo ao primeiro canal direito downmixado e o canal central original. Os segundos canais esquerdo e direito downmixados poderiam então ser transmitidos com todos os quatro conjuntos de códigos de sinalização BCC como informações auxiliares. De forma análoga, um decodificador BCC correspondente poderia então aplicar sequencialmente esses quatro conjuntos de códigos de sinalização BCC em diferentes estágios de um esquema seqüencial de upmixing de dois estágios para sintetizar cinco canais de saída a partir dos dois canais "estéreo" transmitidos.

[00122] Apesar de a presente invenção ter sido descrita no contexto de esquemas de codificação BCC, em que códigos de sinalização ICC combinados são transmitidos com um ou mais canais de áudio (isto é, os E canais transmitidos) juntamente com outros códigos BCC, em configurações alternativas, poderiam ser transmitidos os códigos de sinalização ICC combinados, seja individualmente ou com outros códigos BCC, para um local (ex., um decodificador ou um dispositivo de armazenamento) que já possua os canais transmitidos e possivelmente outros códigos BCC.

[00123] Apesar de a presente invenção ter sido descrita no

contexto de esquemas de codificação BCC, a presente invenção também pode ser implementada no contexto de outros sistemas de processamento de áudio em que os sinais de áudio são decorrelacionados ou outros processamentos de áudio que precisem decorrelacionar sinais.

[00124] Apesar de a presente invenção ter sido descrita no contexto de implementações em que o codificador recebe sinal de áudio de entrada no domínio de tempo e gera sinais de áudio transmitidos no domínio de tempo e o decodificador recebe os sinais de áudio transmitidos no domínio de tempo e gera sinais de áudio playback no domínio de tempo, a presente invenção não é tão limitada. Por exemplo, em outras implementações, qualquer um ou mais dos sinais de áudio de entrada, transmitidos e playback poderiam ser representados em um domínio de frequência.

[00125] Os codificadores e/ou decodificadores BCCs podem ser usados em conjunto ou incorporados a uma variedade de diferentes aplicações ou sistemas, incluindo sistemas para televisão ou de distribuição de música eletrônica, *movie theaters*, difusão, *streaming*, e/ou recepção. Estes incluem sistemas para codificar/decodificar transmissões via, por exemplo, terrestre, por satélite, cabo, internet, intranets ou meios físicos (ex., compact discs, discos digitais versáteis, chips semicondutores, discos rígidos, cartões de memória e similares). Os codificadores e/ou decodificadores BCCs também podem ser empregados em jogos e sistemas de jogos, incluindo, por exemplo, produtos de software interativo para fins de diversão do usuário (ação, role play, estratégia, aventura, simulações, corridas, esportes, fliperama, cartas e jogos de tabuleiro) e/ou para educação, que podem ser publicados para

múltiplas máquinas, plataformas ou mídia. Além disso, os codificadores e/ou decodificadores BCC podem ser incorporados a reprodutores/gravadores de áudio ou sistemas CD-ROM/DVD. Os codificadores e/ou decodificadores BCC também podem ser incorporados a aplicativos de software para PC que incorporam decodificação digital (ex., reprodutor, decodificador) aplicativos de software que incorporam capacidade de codificação digital (ex., codificador, ripador, gravador e jukebox).

[00126] A presente invenção pode ser implementada como processos baseados em circuitos, incluindo possível implementação como circuito integrado simples (como um ASIC ou um FPGA), um módulo multi-chips, cartão simples ou um pacote de circuitos multi-card. Como fica aparente aos técnicos no assunto, podem ser implementadas também várias funções de elementos de circuito, como etapas de processamento em um programa de software. Esse software pode ser empregado, por exemplo, em um processador de sinal digital, micro-controlador ou computador de funções gerais.

[00127] A presente invenção pode ser configurada sob a forma de métodos e equipamentos para a prática desses métodos. A presente invenção também pode ser configurada sob a forma de código de programa configurado em mídia tangível, como disquetes, CD-ROMs, discos rígidos ou qualquer outro meio de armazenamento de leitura por máquina, caracterizado pelo fato de que, quando o código de programas é carregado e executado por uma máquina, como um computador, a máquina se torna um equipamento para a prática da invenção. A presente invenção também pode ser configurada sob a forma de um código de programa, por exemplo, armazenada em um meio de armazenamento, carregada e/ou executada por uma máquina, ou

transmitida por um meio ou portador de transmissão, como uma fiação ou cabeamento elétrico, por meio de fibra ótica ou via radiação eletromagnética, em que, quando o código de programa é carregado e executado por uma máquina, como um computador, a máquina se torna um equipamento para a prática da invenção. Quando implementada em um processador de uso geral, os segmentos do código de programa combinam com o processador para prover um dispositivo exclusivo que opera de forma análoga em circuitos específicos de lógica.

[00128] Será ainda entendido que várias alterações nos detalhes, nos materiais e disposições das peças que foram descritas e ilustradas para explicar a natureza da presente invenção podem ser feitas pelos peritos na técnica sem abandonar o escopo da invenção, como expressa pelas reivindicações a seguir.

[00129] Apesar de as etapas das seguintes reivindicações de método, se houver, serem mencionadas em uma seqüência determinada com a correspondente identificação, a menos que a redação da reivindicação implique de outra forma em uma determinada seqüência para a implementação de alguma ou de todas essas etapas, essas etapas não se destinam necessariamente a ser limitadas para a implementação naquela determinada seqüência.

R E I V I N D I C A Ç Õ E S

1. Método para a codificação de canais de áudio, compreendendo: geração de um ou mais códigos de sinalização **para ser transmitido** para dois ou mais canais de áudio, **caracterizado** pelo fato de que: pelo menos um código de sinalização **dos códigos de sinalização que serão transmitidos** ser um código de sinalização combinado gerado pela combinação de dois ou mais códigos de sinalização estimados **usando uma função média ponderada, a função média ponderada considerando uma importância do canal individual**; e cada código de sinalização estimado é estimado a partir de um grupo diferente de dois ou mais canais de áudio; e transmitir ou armazenar um ou mais códigos de sinalização **para serem transmitidos**.

2. Método, de acordo com a reivindicação 1, **caracterizado** pelo fato de que compreende ainda a transmissão de E canal(is) de áudio transmitido(s) correspondentes aos dois ou mais canais de áudio, em que $E \geq 1$.

3. Método, de acordo com a reivindicação 2, **caracterizado** pelo fato de que: os dois ou mais canais de áudio compreendem C canal(is) de entrada de áudio, em que $C > E$; e os C canais de entrada são downmixados para gerar o(s) E canal(is) transmitido(s).

4. Método, de acordo com a reivindicação 1, **caracterizado** pelo fato de que um ou mais códigos de sinalização **para serem transmitidos** são transmitidos para permitir que um decodificador realize o processamento de síntese durante a decodificação do(s) E canal(is) transmitido(s) com base no código de sinalização combinado, em que o(s) E canal(is) de

áudio transmitido(s) correspondem aos dois ou mais canais de áudio, em que $E \geq 1$.

5. Método, de acordo com a reivindicação 1, **caracterizado** pelo fato de que um ou mais códigos de sinalização **para serem transmitidos** compreendem um ou mais de um código de correlação intercanais (ICC) combinada, um código de diferenças de nível intercanais (ICLD) combinadas e um código de diferenças de tempo intercanais (ICTD) combinadas.

6. Método, de acordo com a reivindicação 1, **caracterizado** pelo fato de que: cada código de sinalização estimado usado para gerar o código de sinalização combinado está associado a um fator ponderado usado na geração de uma média ponderada; e o fator ponderado de cada código de sinalização estimado se baseia na potência do grupo de canais que corresponde ao código de sinalização estimado.

7. Método, de acordo com a reivindicação 1, **caracterizado** pelo fato de que o código de sinalização combinado é um código ICC combinado.

8. Método, de acordo com a reivindicação 7, **caracterizado** pelo fato de que: os dois ou mais canais de áudio compreendem um canal esquerdo, um canal esquerdo traseiro, um canal direito e um canal direito traseiro; é gerado um primeiro código ICC estimado a partir dos canais traseiro e esquerdo traseiro; um segundo código ICC estimado é gerado a partir dos canais direito e traseiro direito; e o código ICC combinado é gerado pela combinação do primeiro e do segundo código ICC estimado.

9. Equipamento para a codificação de canais de

áudio, compreendendo: meios para a geração de um ou mais códigos de sinalização **para serem transmitidos** para dois ou mais canais de áudio, **caracterizado** pelo fato de que: pelo menos um código de sinalização **de um ou mais códigos de sinalização para serem transmitidos** é um código de sinalização combinado gerado pela combinação de dois ou mais códigos de sinalização estimados **usando a função média ponderada, a função média ponderada considerando uma importância do canal individual**; e cada código de sinalização estimado é estimado a partir de um grupo diferente de dois ou mais dos canais de áudio; e meios para a transmissão ou armazenamento de um ou mais códigos de sinalização **para serem transmitidos**.

10. Equipamento, **de acordo com a reivindicação 9, caracterizado pelo equipamento ser** para a codificação de C canais de entrada de áudio para a geração de E canal(is) de áudio transmitido(s), **onde os meios para geração incluem** um estimador de códigos, **sendo que o equipamento adicionalmente compreende** um downmixer adaptado para fazer o downmix de C canais de entrada para a geração de E canal(is) transmitido(s), em que $C > E \geq 1$, **e** pelo fato de que o equipamento é adaptado para transmitir informações **sobre um ou mais** códigos de sinalização **para serem transmitidos**, de maneira a permitir que o decodificador faça o processamento de síntese durante a decodificação do(s) E canal(is) transmitido(s).

11. Equipamento, de acordo com a reivindicação **10, caracterizado** pelo fato de que: o equipamento é um sistema selecionado a partir de um grupo que consiste de um gravador de vídeo digital, um gravador de áudio digital, um computador, um

transmissor por satélite, um transmissor a cabo, um transmissor de difusão terrestre, um sistema de *home entertainment* e um sistema de *movie theater*; e o sistema compreende o estimador de códigos e o downmixer.

12. Fluxo de bits de áudio codificado gerado pela codificação de canais de áudio, **caracterizado** pelo fato de que: um ou mais códigos de sinalização **para serem transmitidos** são gerados para dois ou mais canais de áudio, em que: pelo menos um código de sinalização **dentre os um ou mais códigos de sinalização para serem transmitidos** é um código de sinalização combinado gerado pela combinação de dois ou mais códigos de sinalização estimados **usando a função média ponderada, a função média ponderada considerando uma importância do canal individual**; e cada código de sinalização estimado é estimado a partir de um grupo diferente de dois ou mais dos canais de áudio; e onde os um ou mais códigos de sinalização **para serem transmitidos** e o E canal(is) de áudio transmitido(s) correspondendo aos dois ou mais canais de áudio, em que $E \geq 1$, são codificados no fluxo de bits de áudio codificado.

13. Método para a decodificação de E canal(is) de áudio transmitido(s) para gerar C canais de áudio playback, em que $C > E \geq 1$, o método compreendendo: receber os códigos de sinalização correspondentes ao(s) E canal(is) transmitido(s), **caracterizado** pelo fato de que: pelo menos um código de sinalização **entre os códigos de sinalização recebidos** é um código de sinalização combinado gerado pela combinação de dois ou mais códigos de sinalização estimados; e cada código de sinalização estimado é estimado a partir de um grupo diferente

de duas ou mais **entradas do codificador dos canais de áudio** upmixing de um ou mais do(s) E canal(is) transmitido(s) para gerar um ou mais canais upmixados; e sintetizar um ou mais dos C canais playback aplicando os códigos de sinalização **recebidos** a um ou mais canais upmixados, em que: dois ou mais códigos de sinalização derivados são derivados do código de sinalização combinado; e cada código de sinalização derivado é aplicado para gerar dois ou mais canais sintetizados.

14. Método, de acordo com a reivindicação **13**, **caracterizado** pelo fato de que os códigos de sinalização **recebidos** compreendem um ou mais de um código ICC combinado, um código ICLD combinado e um código ICTD combinado.

15. Método, de acordo com a reivindicação **13**, **caracterizado** pelo fato de que o código de sinalização combinado é uma média de dois ou mais códigos de sinalização estimados em um codificador que recebeu os canais de áudio de entrada do codificador.

16. Método, de acordo com a reivindicação **15**, **caracterizado** pelo fato de que o código de sinalização combinado é uma média ponderada de dois ou mais códigos de sinalização estimados em um codificador que recebeu os canais de áudio de entrada do codificador.

17. Método, de acordo com a reivindicação **16**, **caracterizado** pelo fato de que: cada código de sinalização estimado usado no codificador para gerar o código de sinalização combinado está associado a um fator ponderado usado na geração, no codificador, da média ponderada; e o fator ponderado para cada código de sinalização estimado se baseia em uma potência do

grupo de canais correspondente ao código de sinalização estimado dos canais de áudio de entrada do codificador.

18. Método, de acordo com a reivindicação **13**, **caracterizado** pelo fato de que os dois ou mais códigos de sinalização derivados se derivam por: derivação de um fator de ponderação para cada grupo de dois ou mais canais associados com um código de sinalização estimado; e derivação de dois ou mais códigos de sinalização derivados como uma função do código de sinalização combinado e dois ou mais fatores ponderados derivados.

19. Método, de acordo com a reivindicação **18**, **caracterizado** pelo fato de que cada fator ponderado derivado se deriva por: estimativa de uma potência no grupo de canais correspondentes a um código de sinalização estimado; e derivação do fator ponderado baseado nas potências estimadas de diferentes grupos de canais correspondentes a diferentes códigos de sinalização estimados.

20. Método, de acordo com a reivindicação **13**, **caracterizado** pelo fato de que o código de sinalização combinado é um código ICC combinado.

21. Método, de acordo com a reivindicação **20**, **caracterizado** pelo fato de que: a entrada dos canais de áudio do codificador compreendem um canal esquerdo, um canal traseiro esquerdo, um canal direito e um canal traseiro direito; um primeiro código ICC estimado é gerado em um codificador que recebeu os canais de áudio de entrada do codificador a partir dos canais esquerdo e traseiro esquerdo; um segundo código ICC estimado é gerado no codificador a partir dos canais direito e

traseiro direito; e o código ICC combinado é gerado no codificador pela combinação do primeiro e do segundo códigos ICC estimados.

22. Método, de acordo com a reivindicação **21**, **caracterizado** pelo fato de que: o código ICC combinado é usado, na etapa de sintetizar um ou mais de C canais de playback, para decorrelacionar os canais sintetizados esquerdo e traseiro esquerdo; e o código ICC combinado é usado, na etapa de sintetizar um ou mais de C canais de playback, para decorrelacionar os canais sintetizados direito e traseiro direito.

23. Equipamento para decodificar E canal(is) de áudio transmitido(s) para gerar C canais de áudio playback, em que $C > E \geq 1$, o equipamento compreendendo: meios para recepção dos códigos de sinalização correspondentes aos E canal(is) transmitido(s), **caracterizados** pelo fato de que: pelo menos um código de sinalização **dentre os códigos de sinalização recebidos** é um código de sinalização combinado gerado pela combinação de dois ou mais códigos de sinalização estimados; e cada código de sinalização estimado é estimado a partir de um grupo diferente de **duas ou mais entradas do codificador dos** canais de áudio; meios para upmixing de um ou mais dos E canal(is) transmitido(s) para gerar um ou mais canais upmixados; e meios para sintetizar um ou mais dos C canais playback pela aplicação dos códigos de sinalização **recebidos** a um ou mais canais upmixados; em que: dois ou mais códigos de sinalização derivados são derivados a partir do código de sinalização combinado; e cada código de sinalização derivado é aplicado para gerar dois ou mais canais

sintetizados.

24. Equipamento, de acordo com a reivindicação **23, caracterizado** pelo fato de que: o equipamento é um sistema selecionado a partir de um grupo que consiste de um reproduutor de vídeo digital, um reproduutor de áudio digital, um computador, um receptor por satélite, um receptor a cabo, um receptor de difusão terrestre, um sistema de *home entertainment* e um sistema de *movie theater*; e o sistema compreende o receptor, o upmixer e o sintetizador.

1/9

FIG. 1
(TÉCNICA ANTERIOR)

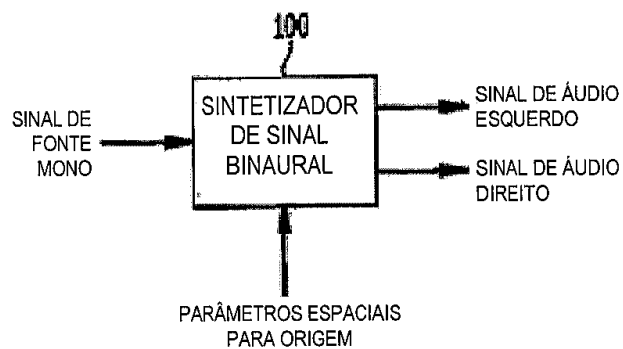


FIG. 2

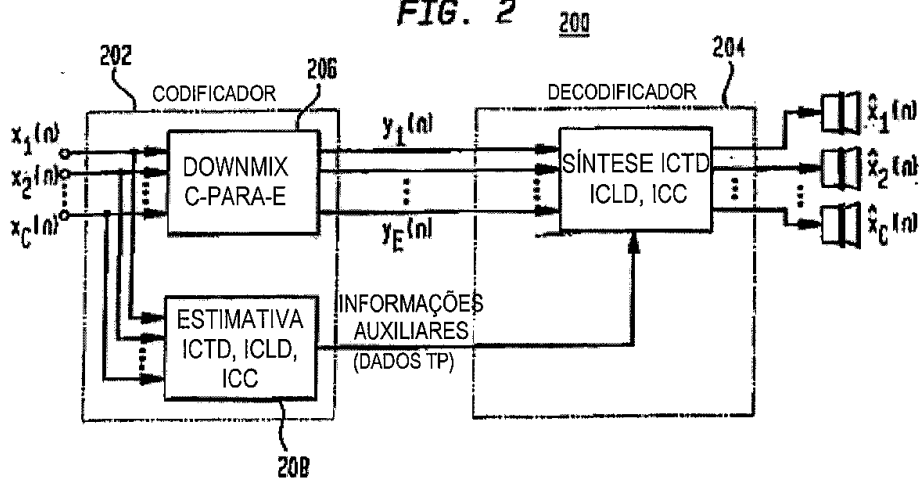


FIG. 3

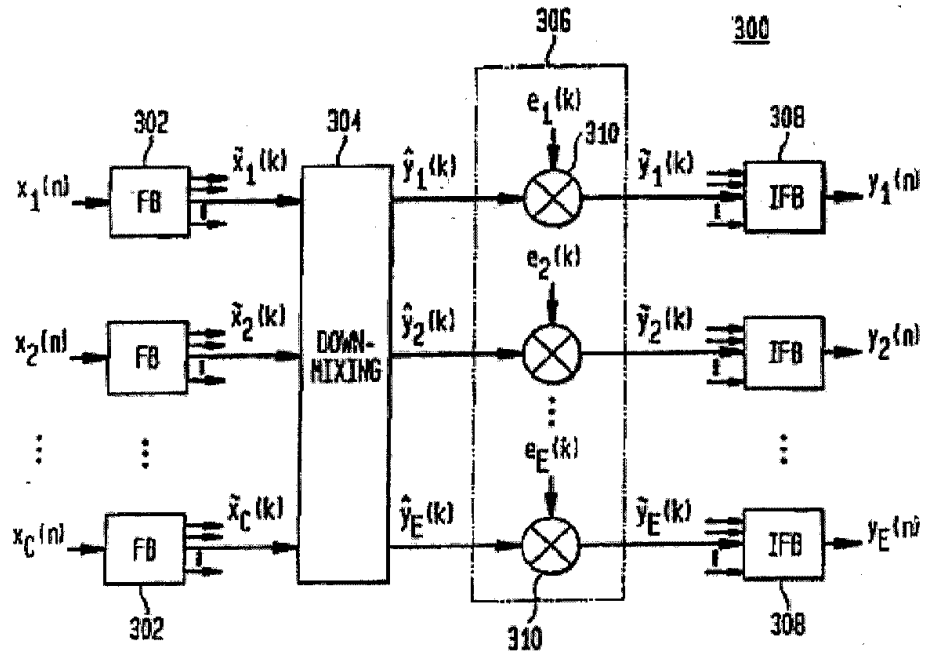
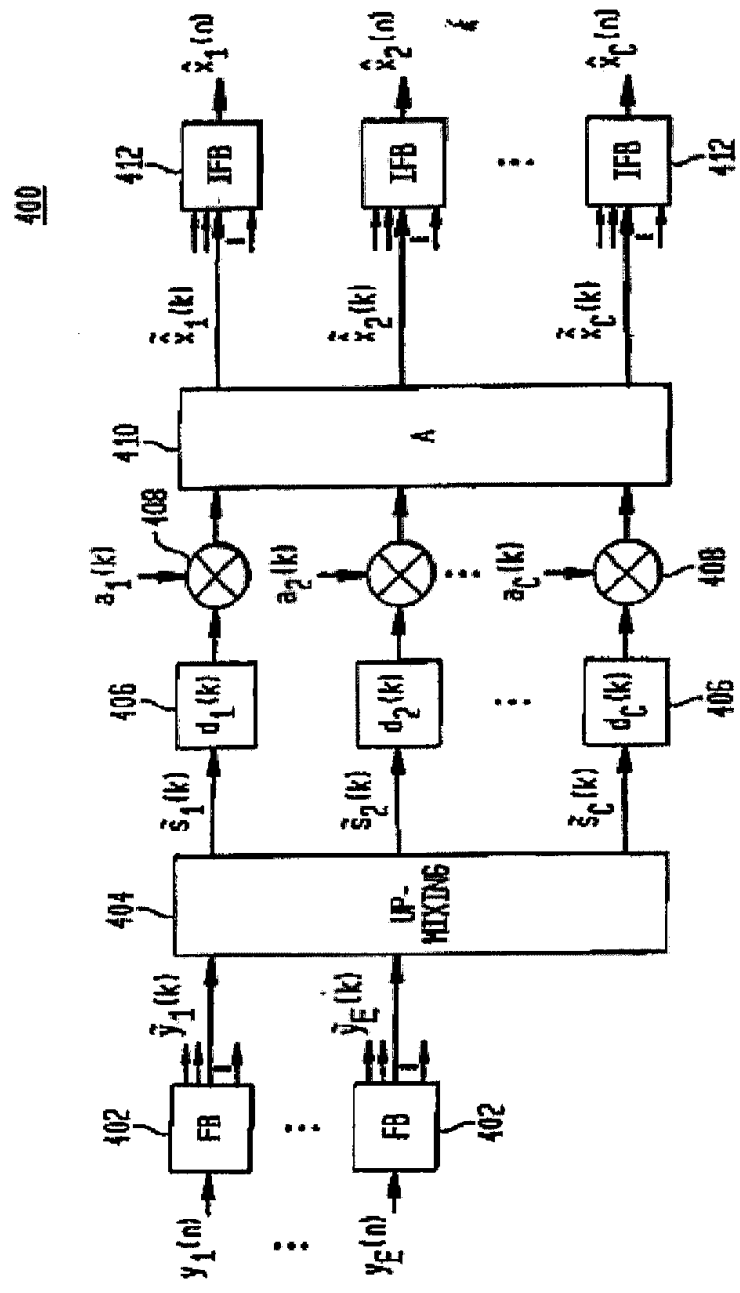


FIG. 4



4/9

FIG. 5

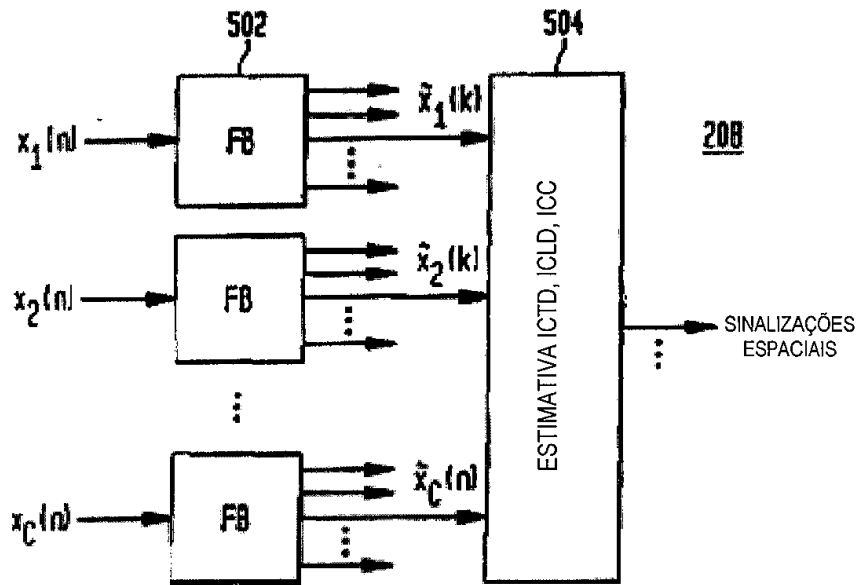
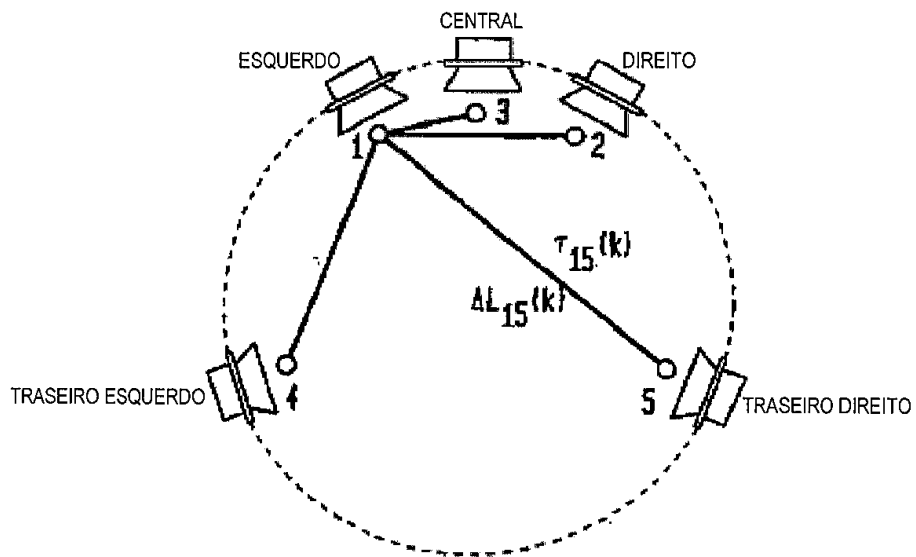


FIG. 6



5/9

FIG. 7A

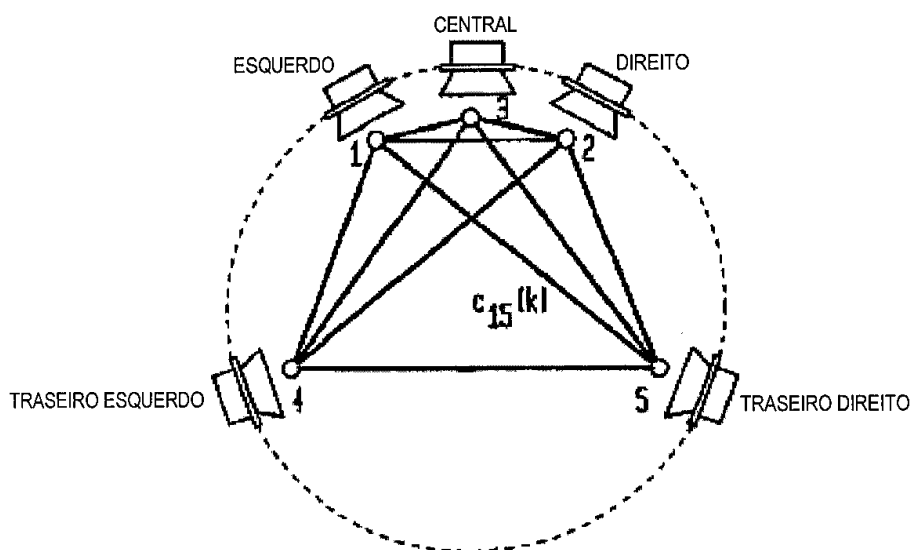


FIG. 7B

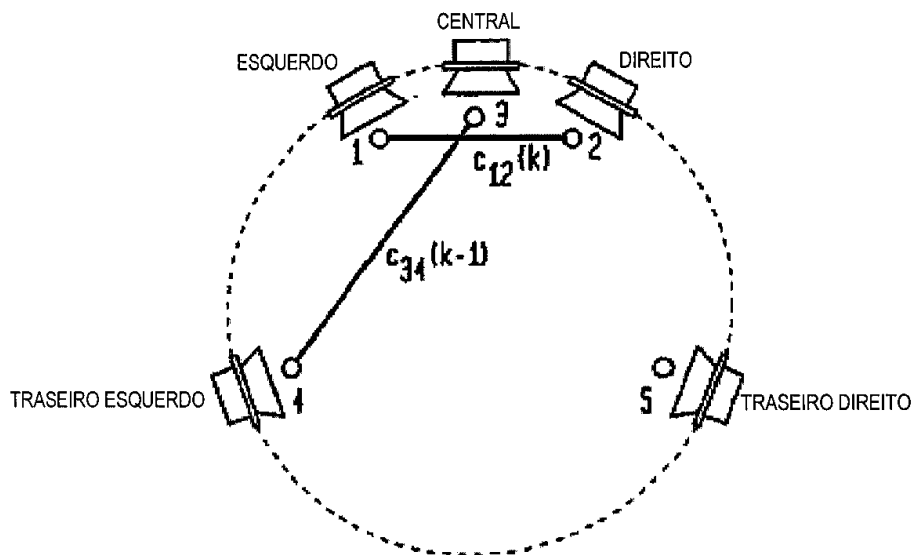
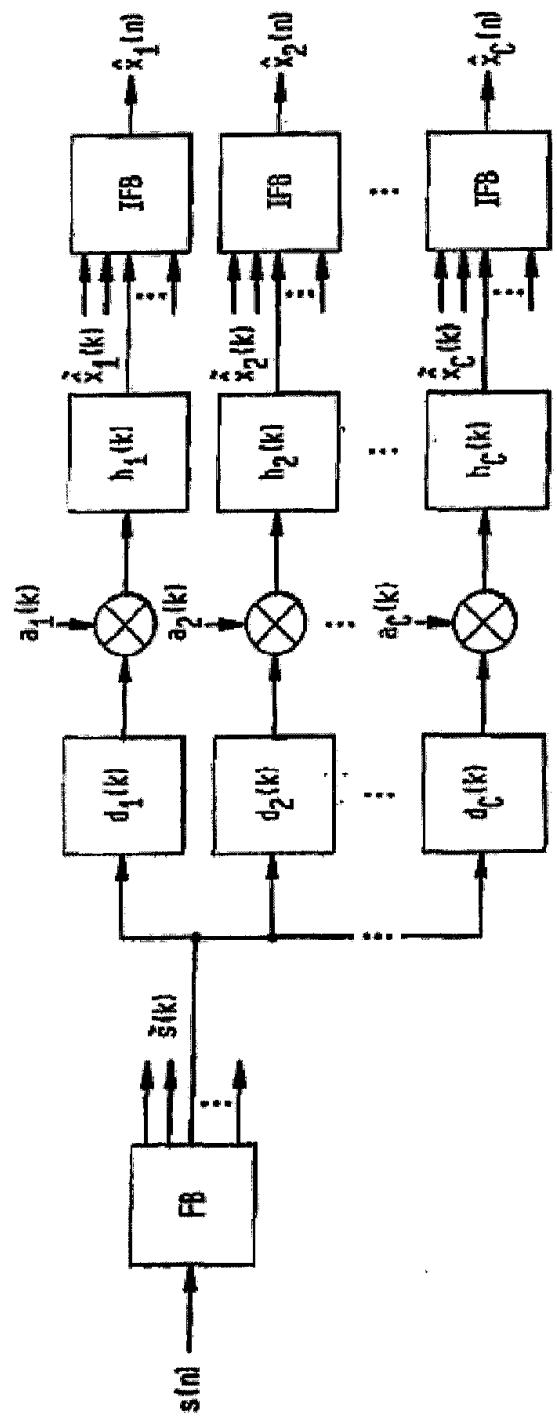


FIG. 8



7/9

FIG. 9

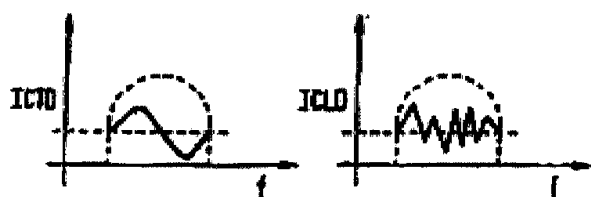


Fig. 10

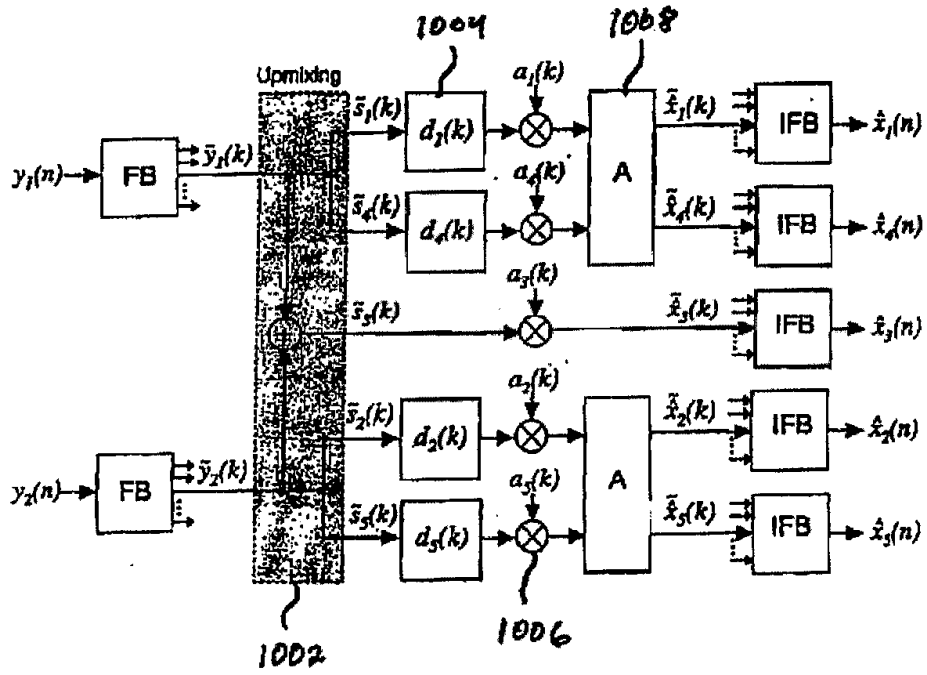


FIG. 11

