



Assinado
Digitalmente

REPÚBLICA FEDERATIVA DO BRASIL
MINISTÉRIO DA INDÚSTRIA, COMÉRCIO EXTERIOR E SERVIÇOS
INSTITUTO NACIONAL DA PROPRIEDADE INDUSTRIAL

CARTA PATENTE Nº PI 0414757-0

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

(21) Número do Depósito: PI 0414757-0

(22) Data do Depósito: 30/09/2004

(43) Data da Publicação do Pedido: 21/04/2005

(51) Classificação Internacional: H04S 3/00.

(30) Prioridade Unionista: US 10/679,085 de 02/10/2003.

(54) Título: EQUIPAMENTO E MÉTODO PARA O PROCESSAMENTO DE UM SINAL DE ÁUDIO MULTICANAIS, EQUIPAMENTO PARA O PROCESSAMENTO INVERSO DOS DADOS DE ENTRADA E MÉTODO DE PROCESSAMENTO INVERSO DOS DADOS DE ENTRADA

(73) Titular: FRAUNHOFER-GESELLSCHAFT ZUR FÖRDERUNG DER ANGEWANDTEN FORSCHUNG E. V., Empresa Alemã. Endereço: Hansastrasse 27 C, 80686 Munich, ALEMANHA(DE)

(72) Inventor: JUERGEN HERRE; JOHANNES HILPERT; STEFAN GEYERSBERGER; ANDREAS HOELZER; CLAUD SPENGER.

Prazo de Validade: 10 (dez) anos contados a partir de 26/12/2018, observadas as condições legais

Expedida em: 26/12/2018

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

MP3

Equipamento e método para o processamento de um sinal de áudio multicanais, equipamento para o processamento inverso dos dados de entrada e método de processamento inverso dos dados de entrada.

5

Campo da invenção

A presente invenção se refere a um aparelho e um método para o processamento de um sinal de áudio multicanais e, em particular, a um aparelho e um método para o processamento de um sinal de áudio multicanais de maneira estéreo-compatível.

10

Histórico da Invenção e da Técnica Anterior

Atualmente, a técnica de reprodução de áudio multicanais está se tornando mais e mais importante. Isto pode ser devido ao fato de que as técnicas de compressão/codificação de áudio como a bem conhecida técnica MP3 tornaram possível distribuir registros de áudio via Internet ou outros canais de transmissão dotados de largura de banda limitada. A técnica de codificação MP3 tornou-se tão famosa devido ao fato de permitir a distribuição de todos os registros em formato estéreo, isto é, uma representação digital do registro de áudio, incluindo um primeiro canal estéreo (ou canal estéreo esquerdo) e um segundo canal estéreo (ou canal estéreo direito).

15

20

Não obstante, existem desvantagens básicas nos sistemas de som convencionais de dois canais. Portanto, foi desenvolvida a técnica surround. Uma representação multicanais surround recomendada inclui, além dos dois canais estéreo L e R, um outro canal central C e dois canais surround Ls, Rs. Esse formato sonoro de referência é também denominado de estéreo três/dois, que significa três canais frontais e dois canais

25

surround. Em geral, são necessários cinco canais de transmissão. Em um ambiente de *playback*, são necessários pelo menos cinco auto-falantes nos respectivos cinco diferentes locais para obter-se um local ideal agradável a uma certa distância a partir dos cinco auto-falantes bem localizados.

São conhecidas várias técnicas nesse campo para reduzir a quantidade de dados necessários para a transmissão de um sinal de áudio multicanais. Essas técnicas são denominadas técnicas *joint stereo*. Para essa finalidade, é feita referência à Fig. 10, que mostra um dispositivo *joint stereo* 60. Esse dispositivo pode ser um dispositivo de implementação, por exemplo, *intensity stereo* (IS) ou *binaural cue coding* (BCC). Esse dispositivo geralmente recebe - como entrada - pelo menos dois canais (CH1, CH2, ... CHn), e emite um único canal portador e dados paramétricos. Os dados paramétricos são definidos de maneira que, em um decodificador, possa ser calculada uma aproximação de um canal original (CH1, CH2, ... CHn).

Normalmente, o canal portador incluirá amostras sub-banda, coeficientes espectrais, amostra de domínio no tempo, etc, que provêm uma representação comparativamente fina do sinal subjacente, enquanto os dados paramétricos não incluem essas amostras de coeficientes espectrais, mas incluem parâmetros de controle para o controle de um determinado algoritmo de reconstrução, como a pesagem por multiplicação, alteração de tempo, alteração de frequência, ... Os dados paramétricos, portanto, incluem somente uma representação comparativamente rústica do sinal ou do canal associado. Explicado em números, a quantidade de dados necessária por um canal portador estará na

faixa de 60 - 70 kbit/s, enquanto a quantidade de dados necessária pelas informações do lado paramétrico para um canal, estará na faixa de 1,5 - 2,5 kbit/s. Um exemplo de dados paramétricos são os bem conhecidos fatores de escala, as informações de *intensity stereo* ou *binaural cue parameters* como será descrito abaixo.

A codificação *intensity stereo* é descrita na pré-impressão AES 3799, "*Intensity Stereo Coding*", J. Herre, K. H. Brandenburg, D. Lederer, February 1994, Amsterdam. Em geral, o conceito de *intensity stereo* se baseia em uma transformada de eixo principal a ser aplicada aos dados de ambos os canais estereofônicos de áudio. Se a maioria dos pontos dos dados está concentrada à volta do primeiro eixo principal, pode ser obtido um ganho de codificação pela rotação de ambos os sinais de um determinado ângulo antes da codificação. Entretanto, isto nem sempre é verdade nas técnicas reais de produção estereofônica. Portanto, essa técnica é modificada pela exclusão da segunda componente ortogonal da transmissão no *bit stream*. Assim, os sinais reconstruídos para os canais esquerdo e direito consistem de versões pesadas ou medidas diferentemente do mesmo sinal transmitido. Não obstante, os sinais reconstruídos diferem em suas amplitudes, mas são idênticos em relação a suas informações de fase. Os envelopes energia-tempo de ambos os canais originais de áudio, entretanto, são preservados por meio de operação de medição seletiva, que tipicamente opera de maneira a selecionar a frequência. Isto se conforma à percepção humana de som em altas frequências, onde os cues espaciais dominantes são determinados pelos envelopes de energia.

Além disso, em implementações práticas, o sinal

transmitido, isto é, o canal portador é gerado a partir do sinal de soma do canal esquerdo e do canal direito ao invés de rotacionar ambos os componentes. Também, esse processamento, isto é, a geração de parâmetros *intensity stereo* para a realização da operação de medição, é feito com seleção de frequência, isto é, independentemente de cada banda de fator de escala, isto é, da partição de frequência codificadora. Preferencialmente, ambos os canais são combinados para formarem um canal combinador ou "portador", e, além do canal combinador, as informações *intensity stereo* determinadas dependem da energia do primeiro canal, da energia do segundo canal ou da energia dos canais combinados.

A técnica BCC é descrita no documento convenção da AES 5574, "*Binaural cue coding applied to stereo and multi-channel audio compression*", C. Faller, F. Baumgarte, May 2002, Munich. Na codificação BCC, alguns canais de entrada de áudio são convertidos em representação espectral usando uma transformada baseada DTF com janelas superpostas. O espectro uniforme resultante se divide em partições não superpostas, cada qual possuindo um índice. Cada partição tem uma largura de banda proporcional à largura de banda retangular equivalente (ERB). As diferenças de nível intercanais (ICLD) e as diferenças de tempos intercanais (ICTD) são estimadas para cada partição para o mesmo quadro *k*. O ICLD e o ICTD são quantificados e codificados, resultando em um *bit stream* BCC. As diferenças de nível intercanais e as diferenças de tempos intercanais são dadas para cada canal com relação a um canal de referência. Então, os parâmetros são calculados de acordo com as formulas indicadas, que dependem de determinadas partições do sinal a ser processado.

No lado decodificador, este recebe um mono-sinal e o *bit stream* BCC. O mono-sinal é transformado em um domínio de frequência e entra em um bloco de síntese espacial, que também recebe valores ICLD e ICTD decodificados. No bloco de síntese espacial, os valores dos parâmetros BCC (ICLD e ICTD) são usados para realizar uma operação de pesagem do mono-sinal para sintetizar os sinais multicanais, que após uma conversão frequência/tempo, representam uma reconstrução do sinal original de áudio multicanais.

No caso BCC, o módulo *joint stereo 60* é operativo para a saída das informações colaterais de canal, de maneira que os dados do canal paramétrico sejam parâmetros ICLD ou ICTD quantificados e codificados, considerando o fato de que um dos canais originais é usado como canal de referência para a codificação das informações colaterais de canal.

Normalmente, o canal portador é formado pela soma dos canais originais participantes.

Naturalmente, as técnicas acima somente provêm uma mono representação para um decodificador, que pode somente processar o canal portador, mas não consegue processar os dados paramétricos para a geração de uma ou mais aproximações de mais de um canal de entrada.

Para a transmissão dos cinco canais de maneira compatível, isto é, em formato *bitstream*, que também é compreensível para um decodificador estéreo normal, a denominada técnica de *matrixing* foi usada como descrito em "*MUSICAM surround: a universal multi-channel coding system compatible with ISO 11172-3*", G. Theile and G. Stoll, AES preprint 3403, October 1992, San

Francisco. Os cinco canais de entrada L, R, C, Ls, e Rs são abastecidos em um dispositivo de *matrixing* e realizam uma operação de *matrixing* para calcular os canais estéreo compatíveis ou básicos Lo, Ro, a partir dos cinco canais de entrada. Em particular, esses canais estéreo básicos Lo/Ro são calculados como

5 indicado abaixo:

$$Lo = L + xC + yLs$$

$$Ro = R + xC + yRs$$

x e y sendo constantes. Os outros três canais C, Ls, Rs são transmitidos como são em uma camada de extensão, além da camada estéreo básica, que inclui uma versão codificada dos sinais estéreo básicos Lo/Ro. Com respeito ao *bitstream*, essa camada estéreo básica Lo/Ro inclui um cabeçalho, informações como os fatores de escala e as amostras de sub-banda. A camada de

10 extensão multicanais, isto é, o canal central e os dois canais *surround* estão incluídos no campo de extensão multicanais, que também é denominado canal de dados auxiliares.

15

No lado decodificador, é realizada uma operação *matrixing* inversa para formar reconstruções dos canais esquerdo e

20 direito na representação de cinco canais usando os canais estéreo básicos Lo, Ro e os três canais adicionais. Além disso, os três canais adicionais são decodificados a partir das informações auxiliares para obter uma representação de cinco canais decodificados ou *surround* do sinal de áudio multicanais original.

25 Outra abordagem de codificação multicanais está descrita na publicação "*Improved MPEG-2 audio multi-channel encoding*", B. Grill, J. Herre, K. H. Brandenburg, E. Eberlein, J. Koller, J. Mueller, AES preprint 3865, February 1994, Amsterdam,

na qual, para se obter uma compatibilidade *backward*, são considerados os modos compatíveis *backward*. Para esse fim, é usada uma matriz de compatibilidade para obter os dois chamados canais *downmix* L_c , R_c a partir dos cinco canais originais de entrada. Além disso, é possível selecionar dinamicamente os três canais auxiliares transmitidos como dados auxiliares.

Para explorar a irrelevância do estéreo, é aplicada uma técnica *joint stereo* aos grupos de canais, por exemplo, os três canais frontais, isto é, para o canal esquerdo, o canal direito e o canal central. Para isso, esses três canais são combinados para obter um canal combinado. Esse canal combinado é quantificado e embalado no *bitstream*. Depois, esse canal combinado juntamente com as correspondentes informações *joint stereo* são colocados em um módulo de decodificação *joint stereo* para obter os canais decodificados *joint stereo* isto é, um canal esquerdo decodificado *joint stereo*, um canal direito decodificado *joint stereo* e um canal central decodificado *joint stereo*. Esses canais decodificados *joint stereo*, juntamente com o canal *surround* esquerdo e o canal *surround* direito entram em um bloco matriz de compatibilidade para formar o primeiro e o segundo canais *downmix* L_c , R_c . Depois, as versões quantificadas do canal combinado são embaladas no *bitstream* juntamente com os parâmetros de codificação *joint stereo*.

Usando codificação *intensity stereo*, portanto, um grupo de sinais independente de canais originais é transmitido dentro de uma porção dos dados "portadores". O decodificador então reconstrói os sinais envolvidos como dados idênticos, que são reescalados de acordo com seus envelopes energia-tempo originais.

Como consequência, uma combinação linear dos canais transmitidos conduzirá aos resultados, que são bastante diferentes do *downmix* original. Isto se aplica a qualquer tipo de codificação *joint stereo* baseada no conceito de *intensity stereo*. Para um sistema de codificação que provê canais *downmix* compatíveis, existe uma consequência direta: A reconstrução por *dematrixing*, como descrita na publicação anterior, sofre por problemas causados pela reconstrução imperfeita. Usando o denominado esquema de pré-distorção *joint stereo*, no qual é feita uma codificação *joint stereo* dos canais da esquerda, da direita e do centro antes de ser feito o *matrixing* no codificador, ameniza o problema. Assim, o esquema de *dematrixing* para a reconstrução introduz menos problemas, já que no lado do codificador, os sinais decodificados *joint stereo* foram usados para gerar os canais *downmix*. Assim, o processo de reconstrução imperfeito é alterado para os canais *downmix* compatíveis L_c e R_c , onde tem muito mais probabilidades de ser mascarado pelo próprio sinal de áudio.

Apesar desse sistema ter resultado em menos problemas devido ao *dematrixing* no lado do decodificador, tem ainda assim algumas desvantagens. A desvantagem é que os canais *downmix* estéreo-compatíveis L_c e R_c se derivam não dos canais originais, mas das versões codificada/decodificada *intensity stereo* dos canais originais. Portanto, as perdas de dados devidas ao sistema de codificação *intensity stereo* estão incluídas nos canais *downmix* compatíveis. O decodificador somente estéreo, que somente decodifica os canais compatíveis ao invés de ampliar os canais codificados *intensity stereo*, portanto, provê um sinal de saída que é afetado pelas perdas de dados induzidas por *intensity*

stereo.

Além disso, um outro canal completo tem que ser transmitido além dos dois canais *downmix*. Esse canal é o canal combinado, que é formado pelos meios de codificação *intensity stereo* do canal esquerdo, do canal direito e do canal central. Além disso, as informações *intensity stereo* para a reconstrução dos canais originais L, R, C do canal combinador também devem ser transmitidas ao decodificador. No decodificador, é feita uma *matrixing* inversa, isto é, é feita uma operação de *dematrixing* para derivar os canais *surround* dos dois canais *downmix*. Também, os canais esquerdo, direito e central originais são aproximados pela decodificação *intensity stereo* usando o canal combinado transmitido e os parâmetros *intensity stereo* transmitidos. Deve ser notado que os canais esquerdo, direito e central originais são derivados pela decodificação *intensity stereo* do canal combinado.

Sumário da Invenção

É o objeto da presente invenção prover um conceito para um processamento bit-eficiente e com a redução de problemas ou o processamento inverso de um sinal de áudio multicanais.

De acordo com um primeiro aspecto da presente invenção, este objeto é alcançado por um equipamento para processamento de um sinal de áudio multicanais, o sinal de áudio multicanais tendo pelo menos três canais originais, compreendendo: meios para prover um primeiro canal *downmix* e um segundo canal *downmix*, o primeiro e o segundo canais *downmix* sendo derivados dos canais originais; meios para calcular as informações colaterais de canal de um canal original selecionado dos sinais originais, os

meios para cálculo sendo operativos para o cálculo das informações colaterais de canal, como de um canal *downmix* ou de um canal *downmix* combinado, incluindo o primeiro e o segundo canais *downmix*, quando pesados usando as informações colaterais de canal, resulta em uma aproximação do canal original selecionado; e meios para a geração de dados de saída, os dados de saída incluindo as informações colaterais de canal, o primeiro canal *downmix* ou um sinal derivado do primeiro canal *downmix* e do segundo canal *downmix* ou um sinal derivado do segundo canal *downmix*.

De acordo com um Segundo aspecto da presente invenção, este objeto é alcançado por um método para o processamento de um sinal de áudio multicanais, o sinal de áudio multicanais tendo pelo menos três canais originais, compreendendo: prover um primeiro canal *downmix* e um segundo canal *downmix*, o primeiro e o Segundo canais *downmix* sendo derivados dos canais originais; calcular as informações colaterais de canal para um canal original selecionado dos sinais originais, de maneira que um canal *downmix* ou um canal *downmix* combinado, incluindo o primeiro e o segundo canais *downmix*, quando pesados usando as informações colaterais de canal, resulta em uma aproximação do canal original selecionado; e gerar dados de saída, os dados de saída incluindo as informações colaterais de canal, o primeiro canal *downmix* ou o sinal derivado do primeiro canal *downmix* e do segundo canal *downmix* ou um sinal derivado do segundo canal *downmix*.

De acordo com um terceiro aspecto da presente invenção, este objeto é alcançado por um equipamento para o processamento inverso dos dados de entrada, os dados de entrada incluindo as informações colaterais de canal, um primeiro canal

downmix ou um sinal derivado do primeiro canal downmix e um
segundo canal downmix ou um sinal derivado do segundo canal
downmix, caracterizado pelo fato de que o primeiro canal downmix e
o segundo canal downmix são derivados de pelo menos três canais
originais de um sinal de áudio multicanais, e caracterizado pelo
fato de que as informações colaterais de canal são calculadas de
maneira que um canal downmix ou um canal downmix combinado,
incluindo o primeiro canal downmix e o segundo canal downmix,
quando pesados usando as informações colaterais de canal, resultam
em uma aproximação do canal original selecionado, o equipamento
compreendendo: uma leitora de dados de entrada para ler os dados
de entrada de maneira a obter o primeiro canal downmix ou um sinal
derivado do primeiro canal downmix e o segundo canal downmix ou um
sinal derivado do segundo canal downmix e as informações
colaterais de canal; e um reconstrutor de canais para a
reconstrução da aproximação do canal original selecionado usando
as informações colaterais de canal e o canal downmix ou o canal
downmix combinado para obter a aproximação do canal original
selecionado.

De acordo com um quarto aspecto da presente
invenção, este objeto é alcançado por um método de processamento
inverso dos dados de entrada, os dados de entrada incluindo as
informações colaterais de canal, um primeiro canal downmix ou um
sinal derivado de um primeiro canal downmix e um segundo canal
downmix ou um sinal derivado de um segundo canal downmix,
caracterizado pelo fato de que o primeiro canal downmix e o
segundo canal downmix são derivados de pelo menos três canais
originais de um sinal de áudio multicanais, e caracterizado pelo

fato de que as informações colaterais de canal são calculadas de maneira que um canal *downmix* ou um canal *downmix* combinado, incluindo o primeiro canal *downmix* e o segundo canal *downmix*, quando pesados usando as informações colaterais de canal, resultam em uma aproximação do canal original selecionado, o método compreendendo: leitura dos dados de entrada para obter o primeiro canal *downmix* ou um sinal derivado do primeiro canal *downmix* e um segundo canal *downmix* ou um sinal derivado do segundo canal *downmix* e as informações colaterais de canal; e reconstruir a aproximação do canal original selecionado usando as informações colaterais de canal e o canal *downmix* ou o canal *downmix* combinado para obter a aproximação do canal original selecionado.

De acordo com um quinto aspecto e sexto aspecto da presente invenção, este objeto é alcançado por um programa de computador incluindo o método de processamento ou o método de processamento inverso.

A presente invenção se baseia na descoberta de que é obtida uma codificação eficiente e com menos problemas do sinal de áudio multicanais quando dois canais *downmix*, preferencialmente representando os canais estéreo esquerdo e direito são embalados nos dados de saída.

Inventivamente, as informações paramétricas colaterais do canal para um ou mais dos canais originais se derivam de maneira a se relacionarem com um dos canais *downmix*, ao invés de, como na técnica anterior, a um canal *joint stereo* "combinado" adicional. Isto significa que as informações paramétricas colaterais do canal são calculadas de maneira que, no lado do decodificador, um reconstrutor de canal usa as informações

colaterais de canal e um dos canais *downmix* ou uma combinação dos canais *downmix* para reconstruir uma aproximação do canal de áudio original, para o qual as informações colaterais de canal são indicadas.

5 O conceito inventivo é vantajoso por prover uma extensão multicanais bit-eficiente, de maneira que possa ser reproduzido em um decodificador o sinal de áudio multicanais.

Além disso, o conceito inventivo é compatível *backward*, já que um decodificador de menor escala, que é somente
10 adaptado para o processamento de dois canais, pode simplesmente ignorar a informação da extensão, isto é, as informações colaterais de canal. O decodificador de menor escala somente pode reproduzir os dois canais *downmix* para obter uma representação estéreo do sinal original de áudio multicanais. Entretanto, um
15 decodificador de maior escala habilitado para operações multicanais, pode usar as informações colaterais de canal transmitidas para reconstruir aproximações dos canais originais.

A presente invenção é vantajosa por ser bit-eficiente, já que, em contraste com a técnica anterior, não é
20 necessário outro canal portador além do primeiro e do segundo canais *downmix* L_c , R_c . Ao invés disso, as informações colaterais de canal se relacionam com um ou ambos os canais *downmix*. Isto significa que os próprios canais *downmix* servem como canais portadores, para os quais as informações colaterais de canal são
25 combinadas para reconstruir um canal de áudio original. Isto significa que as informações colaterais de canal são preferivelmente informações colaterais paramétricas, isto é, informações que não incluem quaisquer amostras de sub-banda ou

coeficientes espectrais. Ao invés disso, as informações colaterais paramétricas são informações usadas para pesagem (no tempo e/ou frequência) do respectivo canal *downmix* ou da combinação dos respectivos canais *downmix* para obter uma versão reconstruída de um canal original selecionado.

Em uma configuração preferida da presente invenção, é obtida uma codificação compatível *backward* de um sinal multicanais baseada em um sinal estéreo compatível. Preferencialmente, o sinal estéreo compatível (sinal *downmix*) é gerado usando o *matrixing* dos canais originais do sinal de áudio multicanais.

Inventivamente, as informações colaterais de canal de um canal original selecionado são obtidas com base nas técnicas *joint stereo* como a codificação *intensity stereo* ou a *binaural cue coding*. Assim, no lado do decodificador, não devem ser feitas operações de *dematrixing*. Os problemas associados com *dematrixing*, isto é, são evitados alguns problemas relacionados com uma distribuição indesejada de ruídos de quantificação nas operações de *dematrixing*. Isto se deve ao fato do decodificador usar um reconstrutor de canais, que reconstrói um sinal original, usando um dos canais *downmix* ou uma combinação dos canais *downmix* e as informações colaterais de canal transmitidas.

Preferivelmente, todo conceito inventivo é aplicado a um sinal de áudio multicanais com cinco canais. Esses cinco canais são um canal esquerdo L, um canal direito R, um canal central C, um canal *surround* esquerdo Ls e um canal *surround* direito Rs. Preferencialmente, os canais *downmix* são canais *downmix* estéreo compatíveis Ls e Rs, que provêm uma representação

estéreo do sinal de áudio multicanais original.

De acordo com a configuração preferida da presente invenção, para cada canal original, são calculadas as informações colaterais de canal em um lado codificador embalado nos dados de saída. As informações colaterais de canal do canal esquerdo original se derivam usando o canal *downmix* esquerdo. As informações colaterais de canal do canal *surround* esquerdo original se derivam usando o canal *downmix* esquerdo. As informações colaterais de canal do canal direito original se derivam do canal *downmix* direito. As informações colaterais de canal para o canal *surround* direito original se derivam do canal *downmix* direito.

De acordo com a configuração preferida da presente invenção, as informações de canal do canal central original se derivam usando o primeiro canal *downmix* assim como o segundo canal *downmix*, isto é, usando uma combinação dos dois canais *downmix*. Preferencialmente, essa combinação é uma soma.

Assim, os subgrupos, isto é, a relação entre as informações colaterais de canal e o sinal portador, isto é, o canal *downmix* usado para prover as informações colaterais de canal para um canal original selecionado é de tal modo que, para a qualidade ideal, é selecionado um determinado canal *downmix*, que contém a maior quantidade relativa possível do respectivo sinal multicanais original que é representado por meio das informações colaterais de canal. Como um sinal portador *joint stereo*, são usados o primeiro e o segundo canais *downmix*. Preferencialmente, também a soma do primeiro e do segundo canais *downmix* pode ser usada. Naturalmente, a soma do primeiro e do segundo canais

downmix pode ser usada para o cálculo das informações colaterais de canal de cada um dos canais originais. Preferencialmente, entretanto, a soma dos canais *downmix* é usada para calcular as informações colaterais de canal do canal central original em um ambiente *surround*, como o *surround* de cinco canais, *surround* de sete canais, *surround* 5.1 ou *surround* 7.1. Usar a soma do primeiro e do segundo canais *downmix* é especialmente vantajoso, já que não precisa ser realizado nenhum outro cabeçalho de transmissão. Isto se deve ao fato de que ambos os canais *downmix* estão presentes no decodificador, de maneira que a soma desses canais *downmix* pode ser facilmente feita no decodificador sem a exigência de outros bits de transmissão.

Preferencialmente, as informações colaterais de canal que formam a extensão multicanais são inseridas no bit stream dos dados de saída de maneira compatível, de maneira que o decodificador de menor escala simplesmente ignora os dados de extensão multicanais e somente fornece uma representação estéreo do sinal de áudio multicanais. Não obstante, um codificador de maior escala não somente usa dois canais *downmix*, mas, além disso, emprega as informações colaterais de canal para reconstruir uma representação completa multicanais do sinal original de áudio.

Um decodificador inventivo é operativo para primeiramente decodificar ambos os canais *downmix* e ler as informações colaterais de canal dos canais originais selecionados. Depois, as informações colaterais de canal e os canais *downmix* são usados para reconstruir aproximações dos canais originais. Para essa finalidade, preferencialmente, não é feita nenhuma operação de *dematrixing*. Isto significa que, nesta configuração, cada um

dos, por exemplo, cinco canais originais de entrada são reconstruídos usando, por exemplo, cinco conjuntos de diferentes informações colaterais de canal. No codificador, é realizado o mesmo agrupamento que no codificador para calcular a aproximação do canal reconstruído. Em um ambiente *surround* de cinco canais, isto significa que, para reconstruir o canal original esquerdo, são usados o canal *downmix* esquerdo e as informações colaterais de canal do canal esquerdo. Para reconstruir o canal direito original, são usados o canal *downmix* direito e as informações colaterais de canal do canal direito. Para reconstruir o canal *surround* esquerdo original, são usados o canal *downmix* esquerdo e as informações colaterais do canal *surround* esquerdo. Para reconstruir o canal *surround* direito original, são usadas as informações colaterais de canal do canal *surround* direito e o canal *downmix* direito. Para reconstruir o canal central original, são usados um canal combinado formado a partir do primeiro canal *downmix* e o segundo canal *downmix* e as informações colaterais do canal central.

Naturalmente, também é possível reproduzir o primeiro e o Segundo canais *downmix* como os canais esquerdo e direito, de maneira que somente três conjuntos (entre, por exemplo, cinco) de parâmetros de informações colaterais do canal tenham que ser transmitidos. Isto é, entretanto, somente aconselhável em situações onde existam regras menos rigorosas com respeito à qualidade. Isto se deve ao fato que, normalmente, o canal *downmix* esquerdo e o canal *downmix* direito são diferentes do canal esquerdo original ou do canal direito original. Somente em situações em que não se possa transmitir as informações colaterais

to canal para cada um dos canais originais, esse processo é vantajoso.

Breve Descrição dos Desenhos

As configurações preferidas da presente invenção são agora discutidas com referência às figuras anexas, nas quais:

Fig. 1 é um diagrama de blocos de uma configuração preferida do decodificador do invento;

Fig. 2 é um diagrama de blocos de uma configuração preferida do decodificador do invento;

10 Fig. 3A é um diagrama de blocos de uma implementação preferida do meio de cálculo para obter informações colaterais de canal seletivo de frequência;

Fig. 3B é uma configuração preferida de um calculador para a implementação de um processamento *joint stereo* como *intensity coding* ou *binaural cue coding*;

Fig. 4 ilustra outra configuração preferida do meio para calcular as informações colaterais de canal, na qual as informações colaterais de canal são fatores de ganho;

Fig. 5 ilustra uma configuração preferida de uma implementação do decodificador, quando o codificador é implementado como na Fig. 4;

Fig. 6 ilustra uma implementação preferida de um meio para prover canais *downmix*;

Fig. 7 ilustra agrupamentos de canais originais e *downmix* para o cálculo das informações colaterais de canal dos respectivos canais originais;

Fig. 8 ilustra outra configuração preferida do codificador do invento;

Fig. 9 ilustra outra implementação de um decodificador do invento; e

Fig. 10 ilustra um codificador *joint stereo* da técnica anterior.

Descrição Detalhada das Configurações Preferidas

A Fig. 1 mostra um equipamento para o processamento de um sinal de áudio multicanais tendo pelo menos três canais originais como R, L e C. Preferencialmente, o sinal de áudio original tem mais do que três canais, como cinco canais no ambiente *surround*, que está ilustrado na Fig. 1. Os cinco canais são o canal esquerdo L, o canal direito R e o canal central C, o canal *surround* esquerdo Ls e o canal *surround* direito Rs. O equipamento do invento inclui o meio 12 para prover um primeiro canal *downmix* Lc e um segundo canal *downmix* Rc, o primeiro e o segundo canais *downmix* sendo derivados dos canais originais. Para derivar os canais *downmix* dos canais originais, existem várias possibilidades. Uma possibilidade é derivar os canais *downmix* Lc e Rc por meio de *matrixing* dos canais originais usando uma operação de *matrixing* como ilustrada na Fig. 6. Essa operação de *matrixing* é realizada no domínio do tempo.

Os parâmetros de *matrixing* a, b e t são selecionados de maneira que sejam menores ou iguais a 1. Preferencialmente, a e b são 0,7 ou 0,5. O parâmetro geral de pesagem t é, preferencialmente, escolhido de maneira que seja evitado o *clipping* de canais.

Alternativamente, como indicado na Fig. 1, os canais *downmix* Lc e Rc podem também ser fornecidos externamente. Isto pode ser feito, quando os canais *downmix* Lc e Rc forem o

resultado de uma operação de "mistura manual". Nesse cenário, o próprio engenheiro de som faz a mixagem dos canais *downmix* ao invés de usar uma operação automatizada de *matrixing*. O engenheiro de som faz uma mixagem criativa para obter os canais *downmix* 5 otimizados L_c e R_c , que proporcionam a melhor representação estéreo possível do sinal de áudio multicanais original.

No caso de um suprimento externo de canais *downmix*, os meios não realizam uma operação *matrixing*, mas simplesmente encaminha os canais *downmix* supridos externamente 10 para um meio subsequente de cálculo 14.

O meio de cálculo 14 é operativo para calcular as informações colaterais de canal como l_i , ls_i , r_i ou rs_i para os canais selecionados originais como L , Ls , R ou Rs , respectivamente. Em particular, o meio 14 para cálculo está 15 operativo para calcular as informações colaterais de canal como de um canal *downmix*, quando pesado usando as informações colaterais de canal, resultando em uma aproximação do canal selecionado original.

Alternativa ou adicionalmente, o meio de cálculo 20 das informações colaterais de canal é ainda operativo e para calcular as informações colaterais de canal para um canal original selecionado como um canal *downmix* combinador, incluindo uma combinação do primeiro e do segundo canais *downmix*, quando pesados usando as informações colaterais de canal calculadas, resulta em 25 uma aproximação do canal original selecionado. Para mostrar essa característica na figura, são mostrados um somador 14a e um calculador de informações colaterais de canal combinado 14b.

Está claro para os peritos na técnica que esses

elementos não precisam ser implementados como elementos distintos. Ao invés disso, toda a funcionalidade dos blocos 14, 14a e 14b pode ser implementada por meio de um determinado processador que pode ser um processador genérico ou qualquer outro meio para a
5 realização da funcionalidade necessária.

Além disso, deve ser notado aqui que os sinais de canal sendo amostras sub-banda ou valores de domínio de frequência estão indicados em letras maiúsculas. As informações colaterais do canal são, em contraste aos próprios canais, indicadas em letras
10 minúsculas. As informações colaterais de canal c_i são, portanto, as informações colaterais de canal do canal central original C.

As informações colaterais de canal, assim como os canais *downmix* L_c e R_c ou uma versão codificada L_c' e R_c' , como produzidas por um codificador de áudio 16 são inseridas em um
15 formatador de dados de saída 18. Em geral, o formatador de dados de saída 18 atua como meio para a geração dos dados de saída, os dados de saída incluindo as informações colaterais de canal de pelo menos um canal original, o primeiro canal *downmix* ou um sinal derivado do primeiro canal *downmix* (como uma versão codificada
20 deste) e o segundo canal *downmix* ou um sinal derivado do segundo canal *downmix* (como uma versão codificada deste).

Os dados de saída ou o *bitstream* de saída 20 pode então ser transmitidos para um decodificador *bitstream* ou pode ser armazenado ou distribuído. Preferencialmente, o *bitstream* de saída
25 20 é um *bitstream* compatível que também pode ser lido por um decodificador de pequena escala não tendo capacidade de extensão multicanais. Esses codificadores de pequena escala, como a maioria dos decodificadores existentes MP3 da técnica atual, simplesmente

ignoram os dados de extensão multicanais, isto é, as informações colaterais de canal. Somente decodificarão o primeiro e o Segundo canais *downmix* para produzir uma saída estéreo. Os decodificadores de maior escala, como os decodificadores habilitados para multicanais lerão as informações colaterais de canal e então gerarão uma aproximação dos canais originais de áudio, como é obtida uma impressão de áudio multicanais.

A Fig. 8 mostra uma configuração preferida da presente invenção no ambiente de cinco canais surround / MP3. Aqui, é preferível escrever os dados de ampliação surround no campo de dados auxiliares na sintaxe *bit stream* MP3 padronizada, como um *bitstream* "MP3 surround" é obtido.

A Fig. 2 mostra uma ilustração de um decodificador do invento atuando como um equipamento de dados de entrada para processamento inverso recebido em uma porta de dados de entrada 22. Os dados recebidos na porta dos dados de entrada 22 são os mesmos dados de saída da porta dos dados de saída 20 da Fig. 1. De maneira alternativa, quando os dados não são transmitidos por canal com fio mas por canal sem fio, os dados recebidos na porta de entrada de dados 22 são derivados dos dados dos dados originais produzidos pelo codificador.

Os dados de entrada do decodificador são inseridos na leitora *data stream* 24 para a leitura dos dados de entrada e finalmente obter as informações colaterais de canal 26 e do canal *downmix* esquerdo 28 e do canal *downmix* direito 30. Caso os dados de entrada incluam versões codificadas dos canais *downmix*, que correspondam ao caso em que o codificador de áudio 16 da Fig. 1 está presente, a leitora *data stream* 24 também inclui um

decodificador de áudio, adaptado ao codificador de áudio usado para a codificação dos canais *downmix*. Nesse caso, o decodificador de áudio, que faz parte da leitora *data stream* 24, é operativo para gerar o primeiro canal *downmix* Lc e o segundo canal *downmix* 5 Rc, ou, melhor explicado, uma versão decodificada desses canais. Para uma descrição mais fácil, a distinção entre sinais e suas versões decodificadas é feita somente quando explicitamente declarada.

As informações colaterais de canal 26 e os canais 10 *downmix* esquerdo e direito 28 e 30 produzidos pela leitora *data stream* 24 são alimentados em um reconstrutor multicanais 32 para prover uma versão reconstruída 34 dos sinais originais de áudio, que podem ser reproduzidos por meio de um *player* multicanais 36. No caso em que o reconstrutor multicanais estiver operativo no 15 domínio da frequência, o *player* multicanais 36 receberá os dados de entrada do domínio de frequência, que devem ser de certa forma decodificados como convertidos no domínio de tempo antes de serem reproduzidos. Para essa finalidade, o *player* multicanais 36 pode também incluir dispositivos para decodificação.

20 Deve ser notado aqui que o decodificador de menor escala somente terá a leitora *data stream* 24, que somente reproduz os canais *downmix* esquerdo e direito 28 e 30 em saída estéreo 38. Entretanto, o decodificador ampliado da invenção extrairá as informações colaterais de canal 26 e usará essas informações 25 colaterais e os canais *downmix* 28 e 30 para reconstruir as versões reconstruídas 34 dos canais originais usando o reconstrutor multicanais 32.

A Fig. 3A mostra uma configuração da calculadora

do invento 14 para o cálculo das informações colaterais de canal que, um codificador de áudio por um lado e a calculadora de informações colaterais de canal pelo outro lado operam na mesma representação espectral do sinal multicanais. Entretanto, a Fig. 1 mostra a outra alternativa, na qual o codificador de áudio por um lado e a calculadora de informações colaterais de canal pelo outro lado operam em diferentes representações espectrais do sinal multicanais. Quando a computação de recursos não é tão importante quanto a qualidade de áudio, é preferida a alternativa da Fig. 1, já que podem ser usados os bancos de filtros otimizados individualmente para codificação de áudio e cálculo de informações colaterais. Entretanto, quando os recursos de computação forem um problema, é preferida a alternativa da Fig. 3A, já que essa alternativa requer menos potência de computação devido à utilização compartilhada dos elementos.

O dispositivo mostrado na Fig. 3A é operativo para a recepção de dois canais A, B. O dispositivo mostrado na Fig. 3A é operativo para calcular as informações colaterais para o canal B, de maneira que usando essas informações colaterais de canal para o canal original selecionado B, pode ser calculada uma versão reconstruída do canal B a partir do sinal do canal A. Além disso, o dispositivo mostrado na Fig. 3A é operativo para formar informações colaterais de canal no domínio de frequência, como parâmetros para a pesagem (multiplicando ou processando no tempo como na codificação BCC, p. exemplo) de valores espectrais ou amostras sub-banda. Para essa finalidade, a calculadora do invento inclui meios de janelamento e de conversão tempo/frequência para obter uma representação de frequência de canal A em uma saída

140b ou uma representação de domínio de frequência do canal B em uma saída 140c.

Na configuração preferida, a determinação das informações colaterais (pelo meio de determinação de informações colaterais 140f) é feita usando valores espectrais quantificados. Depois, também está presente um quantificador 140d que, preferencialmente é controlado usando um modelo psicoacústico tendo uma entrada de controle de modelo psicoacústico 140e. Não obstante, não é necessário um quantificador quando os meios de determinação de informações colaterais 140c usam uma representação não quantificada do canal A para determinar as informações colaterais de canal do canal B.

Caso as informações colaterais de canal do canal B forem calculadas por meio de uma representação de domínio de frequência do canal A e da representação de domínio de frequência do canal B, os meios de janelamento e de conversão tempo/frequência 140a podem ser os mesmos que os usados no codificador de áudio baseado no banco de filtros. Nesse caso, quando é considerada a AAC (ISO/IEC 13818-3), são implementados os meios 140a como um banco de filtros MDCT (MDCT = transformada cosseno discreta modificada) com 50% de funcionalidade *overlap-and-add* [sobreposição e soma].

Nesse caso, o quantificador 140d é um quantificador iterativo como o usado quando são gerados sinais de áudio codificados mp3 ou AAC. A representação do domínio de frequência do canal A que, preferencialmente já está quantificado, pode então ser usada diretamente para a codificação entrópica usando um codificador de entropia 140g, que pode ser um

codificador base Huffman ou um codificador de entropia implementando codificação aritmética.

Quando comparada à Fig. 1, a saída do dispositivo na Fig. 3A são as informações colaterais como a 1, para um canal original (correspondendo às informações colaterais de B na saída do dispositivo 140f). O *bitstream* com codificação entrópica do canal A corresponde, por ex., ao canal codificado *downmix* esquerdo Lc' na saída do bloco 16 da Fig. 1. Na Fig. 3A torna-se claro que o elemento 14 (Fig. 1), i.e., a calculadora para o cálculo das informações colaterais de canal e do codificador de áudio 16 (Fig. 1) pode ser implementado como um meio separado ou pode ser implementado como uma versão compartilhada, de maneira que ambos os dispositivos compartilhem vários elementos como o banco de filtros MDCT 140a, o quantificador 140e e o codificador de entropia 140g. Naturalmente, caso seja necessário uma transformada diferente, etc., para determinar as informações colaterais de canal, então o codificador 16 e a calculadora 14 (Fig. 1) serão implementados em diferentes dispositivos, tais como ambos os elementos não compartilharem o banco de filtros etc.

Em geral, o real determinante para o cálculo das informações colaterais (ou geralmente indicados pela calculadora 14) pode ser implementado como um módulo *joint stereo* como mostrado na Fig. 3B, que opera de acordo com quaisquer das técnicas de *joint stereo* como a codificação *intensity stereo* ou o *binaural cue coding*.

Em contraste com esses codificadores *intensity stereo* da técnica anterior, o meio de determinação do invento 140f não tem que calcular o canal combinado. O "canal combinador" ou

canal portador, como se pode dizer, já existe e é o canal *downmix* compatível esquerdo L_c ou o canal *downmix* compatível direito R_c ou uma versão combinada desses canais *downmix* como $L_c + R_c$. Portanto, o dispositivo do invento 140f somente tem que calcular as medições
5 de escala para escalar o canal *downmix* respectivo, de maneira que seja obtido o envelope energia/tempo do respectivo canal original, quando o canal *downmix* é pesado usando as informações de medição ou, como se pode dizer, as informações direcionais de intensidade.

Portanto, o módulo *joint stereo* 140f da Fig. 3B
10 está ilustrado de maneira a receber, como entrada, o canal A "combinado", que é o primeiro ou o segundo canal *downmix* ou uma combinação dos canais *downmix*, e o canal original selecionado. Esse módulo, naturalmente, reproduz o canal A "combinado" e os parâmetros *joint stereo* como informações colaterais de canal de
15 maneira que, usando o canal A combinado e os parâmetros *joint stereo* possa ser calculada uma aproximação do canal original selecionado.

De maneira alternativa, o módulo *joint stereo* 140f pode ser implementado para realizar o *binaural cue coding*.

20 No caso do BCC, o módulo *joint stereo* 140f é operativo para reproduzir as informações colaterais de canal de maneira que as informações colaterais de canal sejam quantificadas e codificados os parâmetros ICLD ou ICTD, caracterizado pelo fato de que o canal original selecionado serve como o canal real a ser
25 processado, enquanto o canal *downmix* respectivo usado para calcular as informações colaterais, como o primeiro, o segundo ou uma combinação do primeiro e do segundo canais *downmix* seja usada como o canal de referência no sentido da técnica BCC de

codificação/decodificação.

Com referência à Fig. 4, é dada uma simples implementação direcionada para energia do elemento 140f. Esse dispositivo inclui um seletor de banda de frequência 44 que seleciona uma banda de frequência do canal A e uma banda de frequência correspondente do canal B. Depois, em ambas as bandas de frequência, é calculada uma energia por meio de uma calculadora de energia 42 para cada ramal. A implementação detalhada da calculadora de energia 42 dependerá de ser o sinal do bloco 40 é um sinal de sub-banda ou são coeficientes de frequência. Em outras implementações, onde são calculados os fatores de escala para as bandas de fator de escala, pode-se usar já os fatores de escala do primeiro e do segundo canais A, B como valores de energia E_A e E_B ou pelo menos como estimativas de energia. Em um dispositivo de cálculo de fator de ganho 44, é determinado um fator de ganho g_B para a banda selecionada de frequência com base em uma determinada regra, como a regra de determinação de ganho ilustrada no bloco 44 da Fig. 4. Aqui, o fator de ganho g_B pode ser usado diretamente para pesar amostras de domínio de tempo ou coeficientes de frequência, como será descrito posteriormente na Fig. 5. Para esse fim, o fator de ganho g_B , que é válido para a banda selecionada de frequência é usado como informações colaterais de canal para o canal B sendo o canal original selecionado. Esse canal original selecionado B não será transmitido ao decodificador, mas será representado pelas informações colaterais de canal paramétrico calculadas pela calculadora 14 da Fig. 1.

Deve ser aqui notado que não é necessário transmitir valores de ganho como informações colaterais de canal.

É também suficiente transmitir valores dependentes de frequência relativos à energia absoluta do canal original selecionado. Depois, o decodificador deve calcular a energia real do canal *downmix* e o fator de ganho baseado na energia do canal *downmix* e na energia transmitida para o canal B.

A Fig. 5 mostra uma possível implementação de uma montagem de decodificador em conexão com um codificador de áudio perceptual baseado em transformada. Comparado com a Fig. 2, as funcionalidades do decodificador de entropia e do quantificador inverso 50 (Fig. 5) serão incluídas no bloco 24 da Fig. 2. A funcionalidade dos elementos conversores de frequência/tempo 52a, 52b (Fig. 5) será, entretanto implementada no item 36 da Fig. 2. O elemento 50 da Fig. 5 recebe uma versão codificada do primeiro ou do segundo sinal *downmix* Lc' ou Rc' . Na saída do elemento 50, uma versão pelo menos parcialmente decodificada do primeiro e do segundo canais *downmix* está presente, sendo subsequenteiramente denominada de canal A. O canal A é uma entrada em um seletor de banda de frequência 54 para a seleção de uma determinada banda de frequência do canal A. Essa banda selecionada de frequência é pesada usando um multiplicador 56. O multiplicador 56 recebe, para a multiplicação, um determinado fator de ganho g_B , que é indicado para a banda de frequência selecionada pelo seletor de banda de frequência 54, que corresponde ao seletor de banda de frequência 40 da Fig. 4 no lado do codificador. Na entrada do conversor frequência tempo 52a, existe, juntamente com outras bandas, uma representação de domínio de frequência do canal A. Na saída do multiplicador 56 e, em particular na entrada do meio de conversão frequência/tempo 52b será reconstruída a representação de domínio

de frequência do canal B. Portanto, na saída do elemento 52a, será feita uma representação de domínio de tempo para o canal A, enquanto na saída do elemento 52b, haverá uma representação de domínio de tempo do canal B reconstruído.

Deve ser aqui notado que, dependendo de uma determinada implementação, o canal *downmix* decodificado Lc ou Rc não tem *playback* em um decodificador ampliado multicanais. Em tal decodificador ampliado multicanais, os canais *downmix* decodificados são somente usados para reconstruir os canais originais. Os canais *downmix* originais são somente reproduzidos em decodificadores somente estéreo de menor escala.

Para isso, é feita referência à Fig. 9, que mostra a implementação preferida da presente invenção em um ambiente *surround/mp3*. Um *bitstream surround* ampliado mp3 é inserido em um decodificador mp3 padrão 24, que reproduz versões decodificadas dos canais *downmix* originais. Esses canais *downmix* podem então ser diretamente reproduzidos por meio de um decodificador de baixo nível. De maneira alternativa, esses dois canais são inseridos no dispositivo decodificador *joint stereo* avançado 32, que também recebe os dados de extensão multicanais, que são, preferencialmente, inseridos no campo de dados auxiliares em um *bitstream* conforme mp3.

Subseqüentemente, é feita referência à Fig. 7 mostrando o agrupamento do canal original selecionado e o respectivo canal *downmix* ou canal *downmix* combinado. A esse respeito, a coluna direita da tabela da Fig. 7 corresponde ao canal A da Fig. 3A, 3B, 4 e 5, enquanto a coluna do meio corresponde ao canal B nessas figuras. Na coluna esquerda da Fig.

7, as respectivas informações colaterais de canal são explicitamente declaradas. De acordo com a tabela da Fig. 7, as informações colaterais de canal l_i do canal original esquerdo L são calculadas usando o canal *downmix* esquerdo L_c . As informações colaterais do canal *surround* esquerdo ls_i são determinadas por meio do canal *surround* esquerdo original selecionado L_s e o canal *downmix* esquerdo L_c é o portador. As informações colaterais do canal direito r_i do canal direito original R são determinadas usando o canal *downmix* direito R_c . Além disso, as informações colaterais de canal do canal *surround* direito Rs são determinadas usando o canal *downmix* direito R_c como portador. Finalmente, as informações colaterais de canal c_i do canal central C são determinadas usando o canal *downmix* combinado, que são obtidas por meio de uma combinação do primeiro e do segundo canais *downmix*, que podem ser facilmente calculadas tanto em um codificador como em um decodificador e que não exigem bits extra para transmissão.

Naturalmente, também é possível calcular as informações colaterais de canal para os canal esquerdo, por exemplo, baseado em um canal *downmix* combinado ou mesmo em um canal *downmix*, que é obtido por uma adição pesada do primeiro e do segundo canais *downmix*, como $0,7 L_c$ e $0,3 R_c$, enquanto os parâmetros de pesagem são conhecidos de um decodificador ou transmitidos de acordo. Para a maioria das aplicações, entretanto, será preferível somente derivar informações colaterais de canal para o canal central a partir do canal *downmix* combinado, i.e., a partir de uma combinação do primeiro e do segundo canais *downmix*.

Para demonstrar o potencial de economia de bits da presente invenção, é dado o seguinte exemplo típico. No caso de

um sinal de áudio de cinco canais, um codificador comum precisa de uma taxa de bit de 64 kbit/s para cada canal, totalizando uma taxa de bit total de 320 kbit/s para o sinal dos cinco canais. Os sinais estéreo esquerdo e direito requerem uma taxa de bit de 128 kbit/s. As informações colaterais dos canais para um canal ficam entre 1,5 e 2 kbit/s. Portanto, mesmo em um caso em que sejam transmitidas as informações colaterais de canal dos cinco canais, esses dados adicionais somam-se a somente 7,5 a 10 kbit/s. Portanto, o conceito inventivo permite a transmissão de um sinal de áudio de cinco canais usando uma taxa de bit de 138 kbit/s (comparado a 320 (!) kbit/s) com boa qualidade, já que o decodificador não usa a problemática operação de *dematrixing*. Provavelmente ainda mais importante é o fato de que o conceito inventivo é totalmente compatível *backward*, já que cada um dos mp3 *players* pode reproduzir o primeiro canal *downmix* e o segundo canal *downmix* para produzir uma reprodução estéreo convencional.

Dependendo do ambiente da aplicação, o método inventivo para o processamento ou para o processamento inverso pode ser implementado em hardware ou software. A implementação pode ser um meio de armazenagem digital como um disco ou um CD com sinais de controle de leitura eletrônica, que pode cooperar com um sistema de computador programável de maneira que o método do invento para o processamento ou o processamento inverso seja realizado. No geral, portanto, a invenção também se relaciona com um produto de programa de computador dotado de um código de programa armazenado em um portador com leitura por máquina, o código do programa sendo adaptado para realizar o método inventivo, quando o produto de programa de computador opera em um

computador. Em outras palavras, a invenção, portanto, também se relaciona com um programa de computador dotado de um código de programa para a realização do método, quando o programa de computador é operado em um computador.

REIVINDICAÇÕES:

1. "EQUIPAMENTO PARA O PROCESSAMENTO DE UM SINAL DE ÁUDIO MULTICANAIS" tendo, pelo menos, três canais originais, **caracterizado** por:

meios (12) para fornecimento de um primeiro canal de *downmix* como um canal de *downmix* esquerdo e um segundo canal de *downmix* como um canal de *downmix* direito, o primeiro e segundo canais de *downmix* sendo derivados dos canais originais, de modo que os canais de *downmix* esquerdo e direito sejam formados tal que o resultado, quando reproduzido, seja uma representação estéreo do sinal de áudio multicanal;

meios (14) para cálculo da informação secundária para um canal original selecionado dos sinais originais, o meio para cálculo sendo operativo para calcular a informação secundária do canal, de modo que um canal de *downmix* ou um canal de *downmix* combinado, incluindo o primeiro e segundo canais de *downmix*, quando ponderados utilizando a informação secundária do canal, resulte em uma aproximação do canal original selecionado; e

meios (18) para geração de dados de saída, os dados de saída incluindo a informação secundária do canal.

2. "EQUIPAMENTO" de acordo com a reivindicação 1, **caracterizado** pelos meios (18) para geração serem operativos para gerar os dados de saída, de modo que os dados de saída adicionalmente incluam o primeiro canal de *downmix* ou um sinal derivado do primeiro canal de *downmix* e o segundo canal de *downmix* ou um sinal derivado do segundo canal de

downmix.

3. "EQUIPAMENTO" de acordo com a reivindicação 1 ou 2, **caracterizado** pelos meios (14) para cálculo serem operativos para determinar a informação secundária do canal como dados paramétricos, não incluindo amostras de domínio de tempo ou valores espectrais.

4. "EQUIPAMENTO" de acordo com qualquer uma das reivindicações de 1 a 3, **caracterizado** pelos meios (14) para cálculo serem operativos para realizar a codificação estéreo conjunta, utilizando um canal de *downmix* como um canal condutor e utilizando, como um canal de entrada, o canal original selecionado para gerar parâmetros estéreo conjuntos como informação secundária do canal para o canal original selecionado.

5. "EQUIPAMENTO" de acordo com a reivindicação 3, na qual os meios (14) para cálculo são operativos para realizar codificação estéreo de intensidade ou codificação de sinal binaural, de modo que a informação secundária do canal represente uma distribuição de energia ou parâmetros de sinal binaural para o canal original selecionado, **caracterizado** pelo canal de *downmix* ou um canal de *downmix* combinado ser utilizável como um canal condutor.

6. "EQUIPAMENTO" de acordo com as reivindicações de 1 a 5, **caracterizado** pelo fato de que o sinal de áudio multicanal inclui um canal esquerdo, um canal surround esquerdo, um canal direito e um canal surround direito;

nas quais os meios (12) para fornecimento são operativos para fornecer o primeiro canal de *downmix* como um

canal de *downmix* esquerdo e para fornecer o segundo canal de *downmix* como um canal de *downmix* direito, os canais de *downmix* esquerdo e direito sendo formados tal que um resultado, quando reproduzido, seja uma representação estéreo do sinal de áudio multicanal; e

nas quais os meios (14) para cálculo são operativos:

para calcular a informação secundária do canal para o canal esquerdo como o canal original selecionado, utilizando o canal de *downmix* esquerdo;

para calcular a informação secundária do canal para o canal direito como o canal original selecionado, utilizando o canal de *downmix* direito;

para calcular a informação secundária do canal para o canal surround esquerdo como o canal original selecionado utilizando o canal de *downmix* esquerdo; e

para calcular a informação secundária do canal para o canal surround direito como o canal original selecionado utilizando o canal de *downmix* direito.

7. "EQUIPAMENTO" de acordo com as reivindicações de 1 a 6,

nas quais os canais originais incluem um canal central que inclui, ainda, um combinador (14a) para combinar o primeiro canal de *downmix* e o segundo canal de *downmix* para obter o canal de *downmix* combinado; e

caracterizado pelos meios para cálculo da informação secundária do canal para o canal central como o canal original selecionado serem operativos para calcular (14b)

a informação secundária do canal, de modo que o canal de *downmix* combinado, quando ponderado utilizando a informação secundária do canal, resulte em uma aproximação do original canal central.

8. "EQUIPAMENTO" de acordo com qualquer uma das reivindicações de 1 a 6, **caracterizado** pelos meios (12) para fornecimento serem operativos para derivar o primeiro canal de *downmix* e o segundo canal de *downmix* dos canais originais, utilizando uma primeira combinação ponderada linear predeterminada para o primeiro canal de *downmix* e utilizando uma segunda combinação ponderada linear predeterminada para o segundo canal de *downmix*.

9. "EQUIPAMENTO" de acordo com a reivindicação 7, na qual a primeira combinação ponderada linear predeterminada é definida, conforme segue:

$$L_c = t \cdot (L + a \cdot L_s + b \cdot C); \text{ ou}$$

na qual a segunda combinação ponderada linear predeterminada é definida, conforme segue:

$$R_c = t \cdot (R + a \cdot R_s + b \cdot C),$$

caracterizado por L_c ser o primeiro canal de *downmix*, em que R_c é o segundo canal de *downmix*, em que t , a e b são fatores de ponderação menores do que 1, em que L é um canal esquerdo original, em que C é um canal central original, em que R é um canal direito original, em que L_s é um canal surround esquerdo original e em que R_s é um canal surround direito original.

10. "EQUIPAMENTO" de acordo com qualquer uma das reivindicações de 1 a 8, **caracterizado** pelos meios (12)

para fornecimento serem operativos para receber primeiro e segundo canais de *downmix* externamente fornecidos.

11. "EQUIPAMENTO" de acordo com qualquer uma das reivindicações de 1 a 10, **caracterizado** pelo primeiro canal de *downmix* e pelo segundo canal de *downmix* serem canais compostos, sendo compostos dos canais originais em graus variáveis, em que os meios para cálculo são operativos, em uso, para calcular a informação secundária do canal, o canal de *downmix* entre ambos canais de *downmix*, o que é mais fortemente influenciado pelo canal original selecionado quando comparado com outro canal de *downmix*.

12. "EQUIPAMENTO" de acordo com qualquer uma das reivindicações de 1 a 11, **caracterizado** pelos meios (18) para gerar ser operativo para formar os dados de saída de modo que os dados de saída estejam de acordo com uma sintaxe dos dados de saída a ser utilizada por um decodificador de baixo nível para processamento do primeiro canal de *downmix* ou um sinal derivado do primeiro canal de *downmix* ou do segundo canal de *downmix* ou um sinal derivado do segundo canal de *downmix* para obter uma representação estéreo decodificada do sinal de áudio multicanal.

13. "EQUIPAMENTO" de acordo com a reivindicação 12, **caracterizado** pela sintaxe dos dados de saída é estruturada, de modo que a mesma inclua um campo de dados especiais a ser ignorado por um decodificador de baixo nível e na qual os meios para geração são operativos para inserir a informação secundária do canal no campo de dados especiais.

14. "EQUIPAMENTO" de acordo com a reivindica-

ção 13, **caracterizado** pela sintaxe ser uma sintaxe de mp3 e o campo de dados especiais é um campo de dados auxiliar.

15. "EQUIPAMENTO" de acordo com qualquer uma das reivindicações de 12 a 14, **caracterizado** pelos meios (18) para geração são operativos para inserir a informação secundária do canal nos dados de saída, de modo que a informação secundária do canal seja apenas utilizada por um decodificador de alto nível, mas seja ignorada pelo decodificador de baixo nível.

16. "EQUIPAMENTO" de acordo com qualquer uma das reivindicações de 2 a 15, **caracterizado** por compreender, ainda, um codificador (16) para codificação do primeiro canal de *downmix* para obter o sinal derivado do primeiro canal de *downmix* ou para codificação do segundo canal de *downmix* para obter o sinal derivado do segundo canal de *downmix*.

17. "EQUIPAMENTO" de acordo com a reivindicação 16, **caracterizado** pelo codificador (16) ser um codificador perceptual que inclui meios para conversão de um sinal a ser codificado em uma representação espiral, meio para quantização da representação espiral, utilizando um modelo psicoacústico, e meios para codificação por entropia de uma representação espiral quantizada para obter uma representação espiral quantizada codificada por entropia como o sinal derivado do primeiro canal de *downmix* ou o sinal derivado do segundo canal de *downmix*.

18. "EQUIPAMENTO" de acordo com a reivindicação 17, **caracterizado** pelo codificador perceptual (16) ser um codificador de acordo com a camada III (mp3) de MPEG-1/2 ou

codificação de áudio avançada (AAC) de MPEG-2/4.

19. "EQUIPAMENTO" de acordo com qualquer uma das reivindicações de 1 a 18, **caracterizado** pelos meios (14) para cálculo serem operativos para calcular valores de energia de *downmix* para o canal de *downmix* ou o canal de *downmix* combinado;

para calcular um valor de energia original para o canal original selecionado; e

para calcular um fator de ganho como a informação secundária do canal, o fator de ganho sendo derivado do valor de energia de *downmix* e o valor de energia original.

20. "EQUIPAMENTO" de acordo com qualquer uma das reivindicações de 1 a 19, **caracterizado** pelos meios (14) para cálculo serem operativos para calcular parâmetros de informação secundária do canal dependente de frequência, de modo que, para uma pluralidade de bandas de frequência, uma pluralidade de diferentes parâmetros de informação secundária do canal seja obtida.

21. "MÉTODO PARA O PROCESSAMENTO DE UM SINAL DE ÁUDIO MULTICANAIS" tendo, pelo menos, três canais originais, **caracterizado** por compreender:

fornecimento (12) de um primeiro canal de *downmix* como um canal de *downmix* esquerdo e um segundo canal de *downmix* como um canal de *downmix* direito, o primeiro e segundo canais de *downmix* sendo derivados dos canais originais, de modo que os canais de *downmix* esquerdo e direito resultem, quando reproduzidos, em uma representação estéreo do sinal de áudio multicanal;

cálculo (14) da informação secundária do canal para um canal original selecionado dos sinais originais, de modo que um canal de *downmix* ou um canal de *downmix* combinado, incluindo o primeiro e segundo canais de *downmix*, quando ponderados utilizando a informação secundária do canal, resulte em uma aproximação do canal original selecionado; e

geração (18) de dados de saída, os dados de saída incluindo a informação secundária do canal.

22. "EQUIPAMENTO PARA O PROCESSAMENTO INVERSO DOS DADOS DE ENTRADA", os dados de entrada compreendendo informação do lado do canal, um canal de *downmix* esquerdo ou um sinal derivado do canal de *downmix* esquerdo e um canal de *downmix* direito ou um sinal derivado do canal de *downmix* direito, sendo que pelo canal de *downmix* esquerdo e pelo canal de *downmix* direito serem derivados de, pelo menos, três canais originais de um sinal de áudio multicanal e resultarem, quando reproduzidos, em uma representação estéreo do sinal de áudio multicanal e em que a informação secundária do canal é calculada, de modo que um canal de *downmix* ou um canal de *downmix* combinado, incluindo o canal de *downmix* esquerdo e o canal de *downmix* direito, quando ponderados utilizando a informação secundária do canal, resulte em uma aproximação do canal original selecionado, o aparelho compreendendo:

um leitor de dados de entrada (24) para leitura dos dados de entrada para obter o canal de *downmix* esquerdo ou um sinal derivado do canal de *downmix* esquerdo e o canal de *downmix* direito ou um sinal derivado do canal de *downmix* direito e a informação secundária do canal; e

um reconstrutor de canal (32) para reconstruir a aproximação do canal original selecionado, utilizando a informação secundária do canal e do canal de *downmix* ou do canal de *downmix* combinado para obter a aproximação do canal original selecionado;

caracterizado pelo reconstrutor de canal ser operativo para reconstruir uma aproximação para o canal central, usando informações colaterais de canal para o canal central e para o canal de *downmix* combinado.

23. "EQUIPAMENTO" de acordo com a reivindicação 22, **caracterizado** por compreender, ainda, um decodificador perceptual (24) para decodificar o sinal derivado do canal de *downmix* esquerdo para obter a versão decodificada do canal de *downmix* esquerdo e para decodificar o sinal derivado do canal de *downmix* direito para obter uma versão decodificada do canal de *downmix* direito.

24. "EQUIPAMENTO" de acordo com a reivindicação 22 ou 23, **caracterizado** por compreender, ainda, um combinador para combinar o canal de *downmix* esquerdo e o canal de *downmix* direito para obter o canal de *downmix* combinado.

25. "EQUIPAMENTO" de acordo com a reivindicação 22, **caracterizado** pelo aparelho ser configurado para não realizar qualquer operação de desmatrização.

26. "MÉTODO DE PROCESSAMENTO INVERSO DOS DADOS DE ENTRADA", incluindo informação secundária do canal, um canal de *downmix* esquerdo ou um sinal derivado do canal de *downmix* esquerdo e um canal de *downmix* direito ou um sinal derivado do canal de *downmix* direito, sendo que pelo canal de

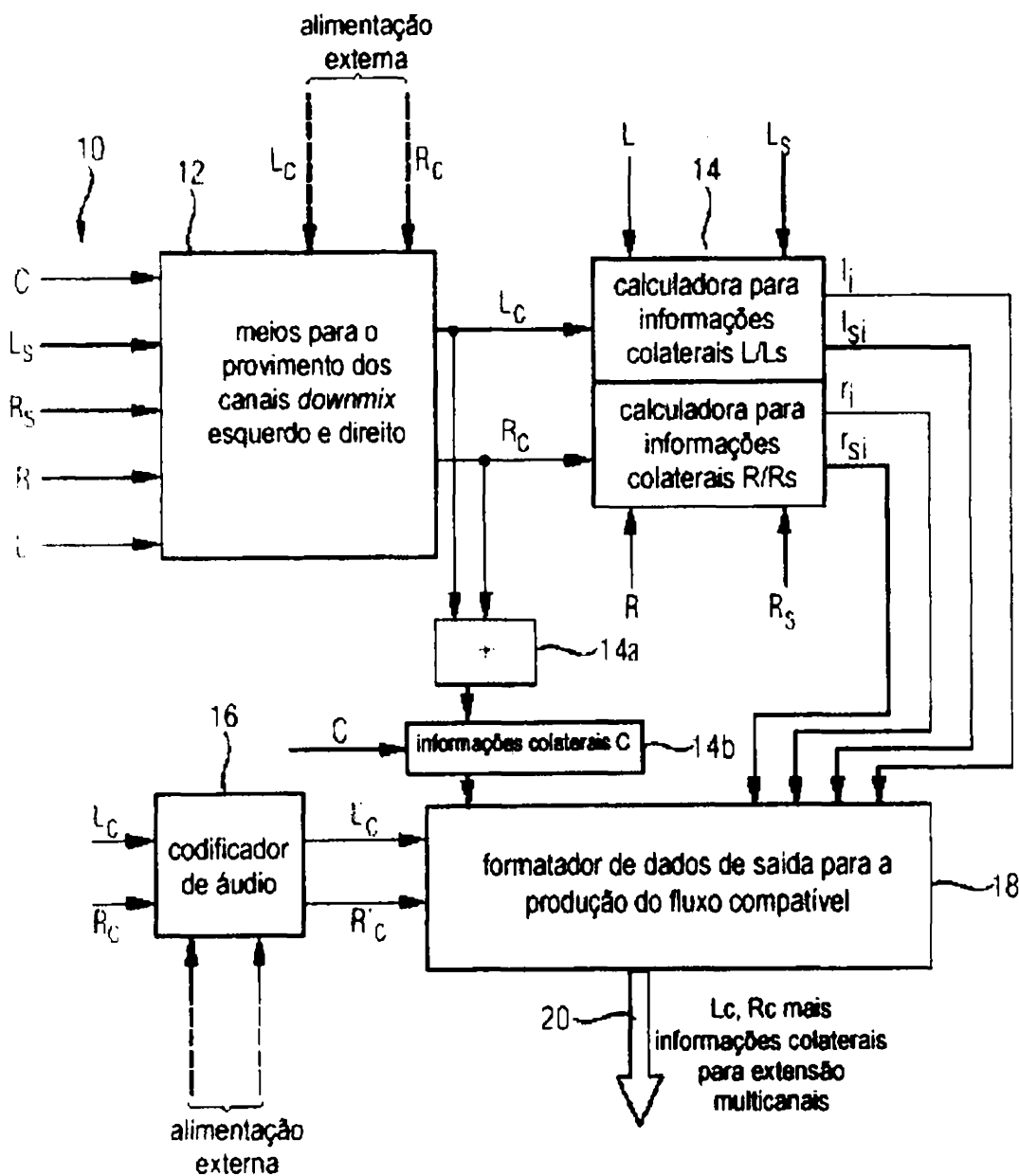
downmix esquerdo e o canal de *downmix* direito serem derivados de, pelo menos, três canais originais de um sinal de áudio multicanal e resultarem, quando reproduzidos, em uma representação estéreo do sinal de áudio multicanal e em que a informação secundária do canal é calculada, de modo que um canal de *downmix* ou um canal de *downmix* combinado, incluindo o canal de *downmix* esquerdo e o canal de *downmix* direito, quando ponderados utilizando a informação secundária do canal, resulte em uma aproximação do canal original selecionado, o método compreendendo:

leitura (24) dos dados de entrada para obter o canal de *downmix* esquerdo ou um sinal derivado do canal de *downmix* esquerdo e o canal de *downmix* direito ou um sinal derivado do canal de *downmix* direito e a informação secundária do canal; e

reconstrução (32) da aproximação do canal original selecionado, utilizando a informação secundária do canal e do canal de *downmix* ou do canal de *downmix* combinado para obter a aproximação do canal original selecionado;

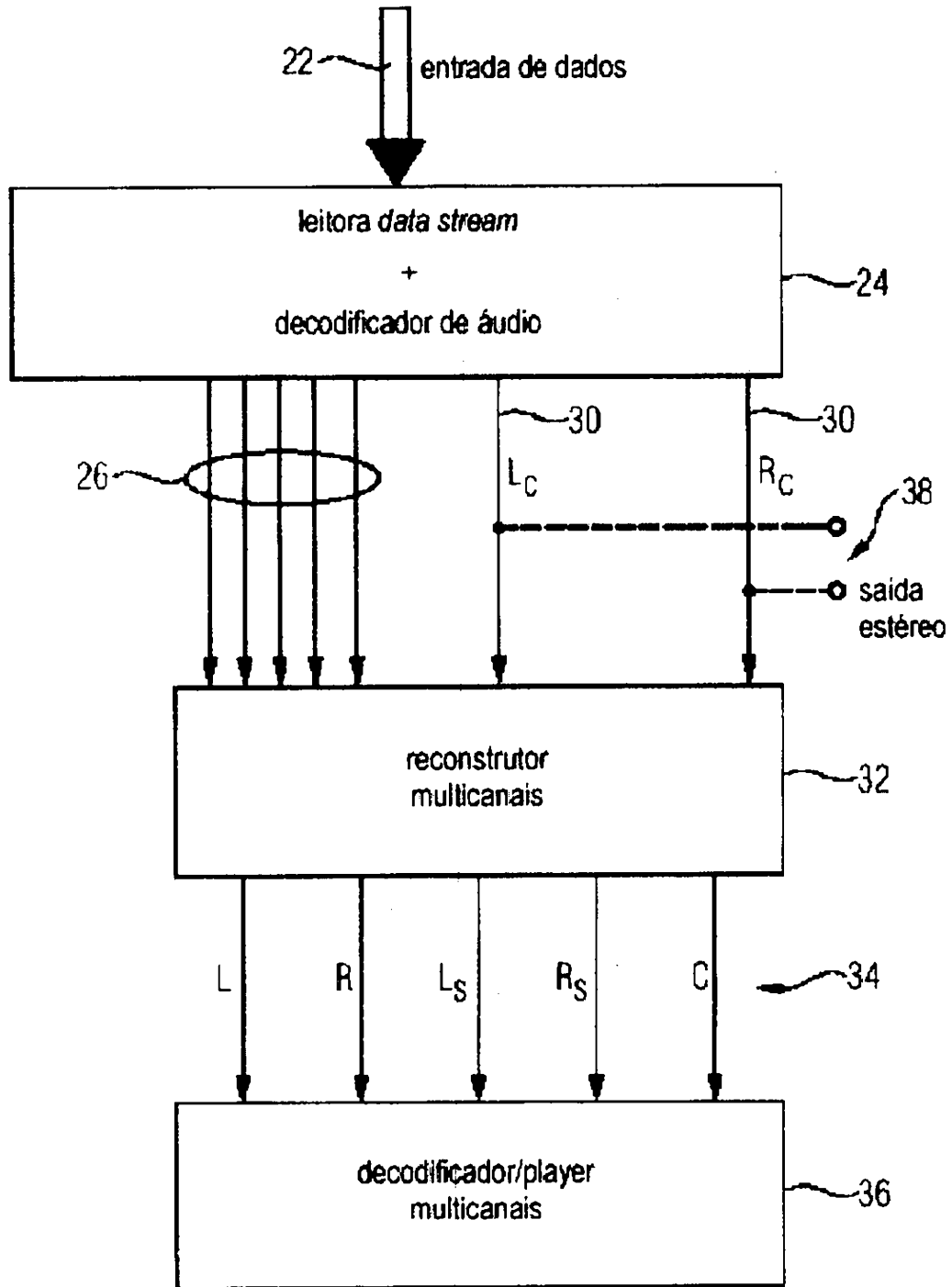
caracterizado pela reconstrução compreender reconstrução de uma aproximação para o canal central usando informações colaterais de canal para o canal central e o canal de *downmix* combinado.

FIG 1 (Codificador)



2/7

FIG 2 (Decodificador)



3/7

FIG 3A

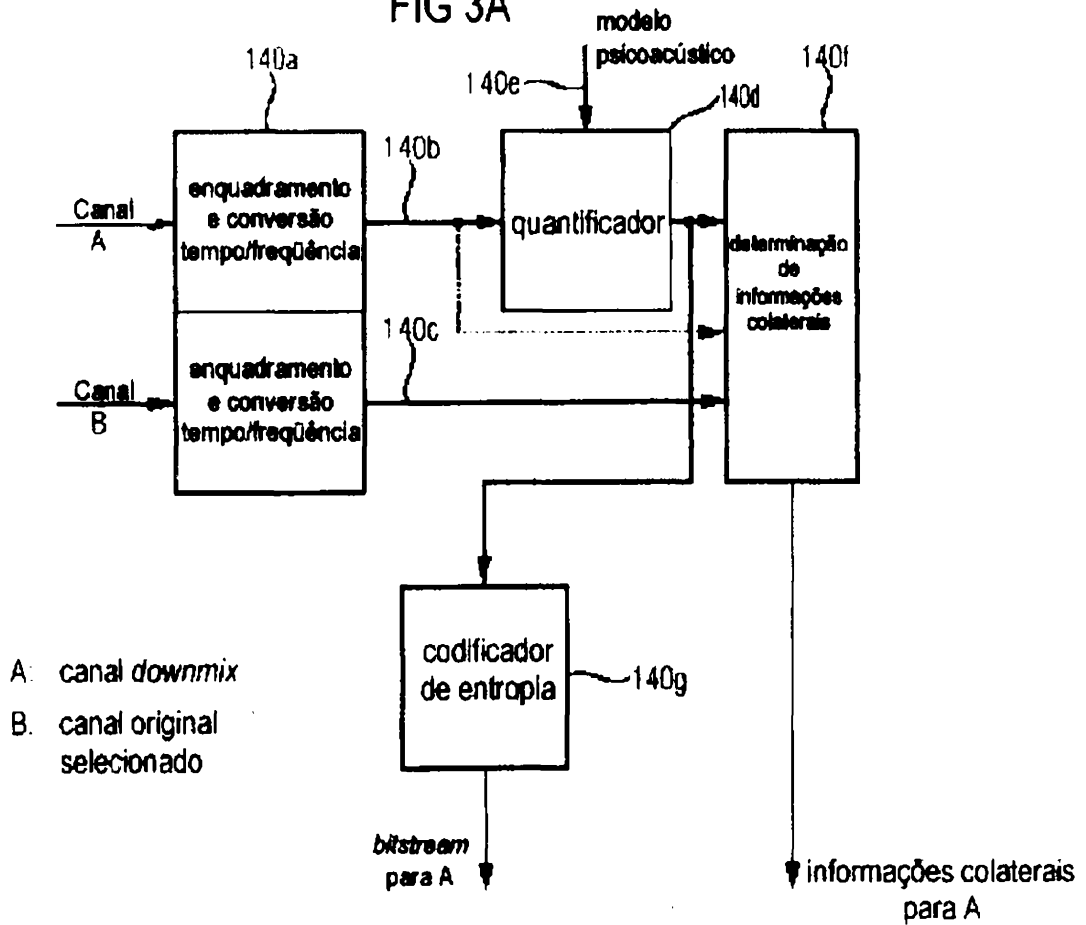
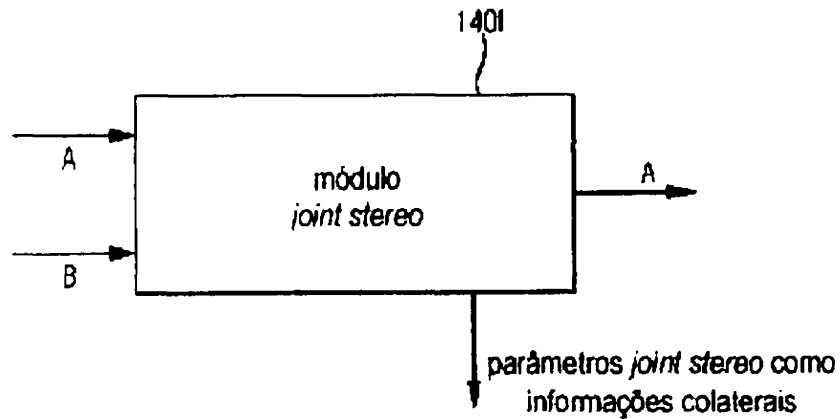


FIG 3B



4/7

FIG 4 (lado codificador)

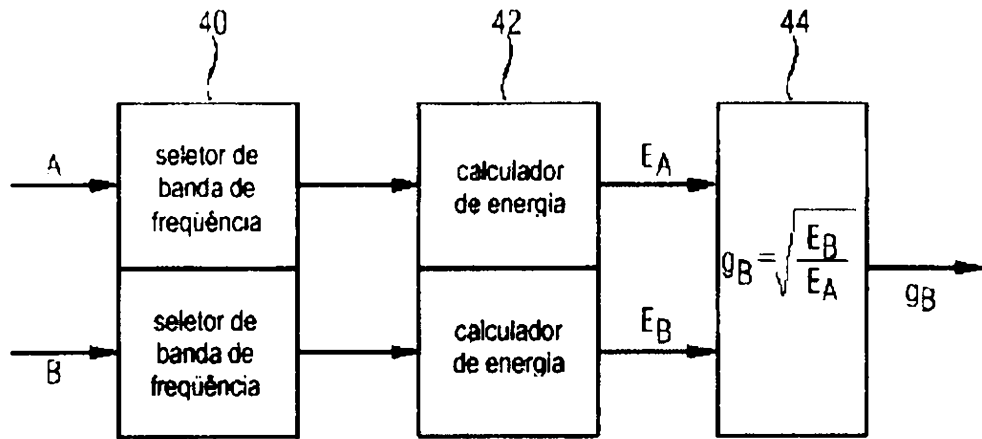
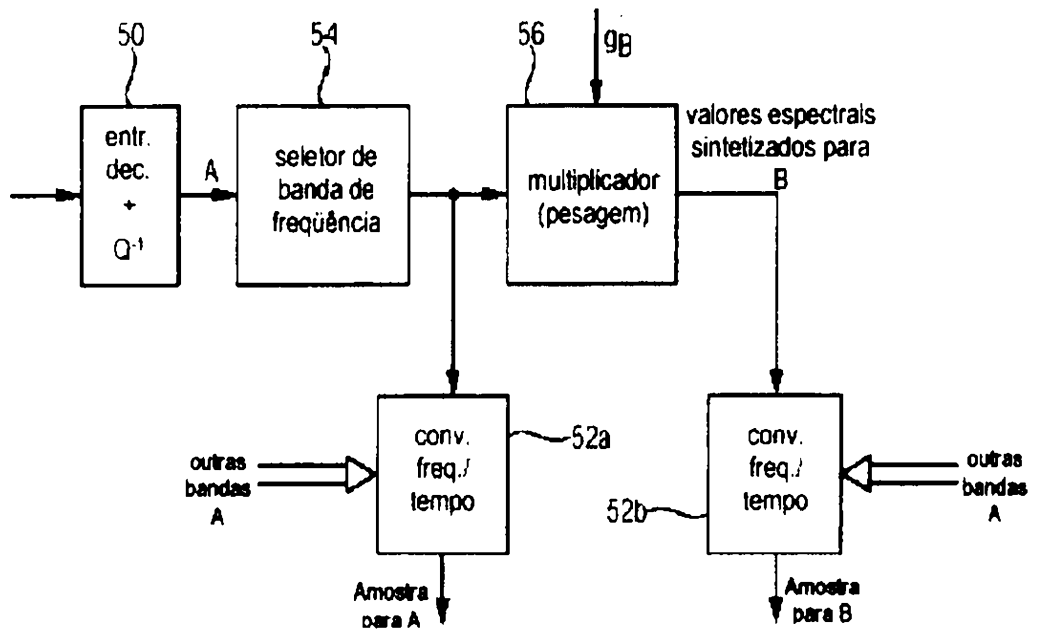


FIG 5 (lado decodificador)



5/7

FIG 6

DOWNMIX

$L_C = t \cdot [L + a \cdot L_S + b \cdot C]$
$R_C = t \cdot [R + a \cdot R_S + b \cdot C]$

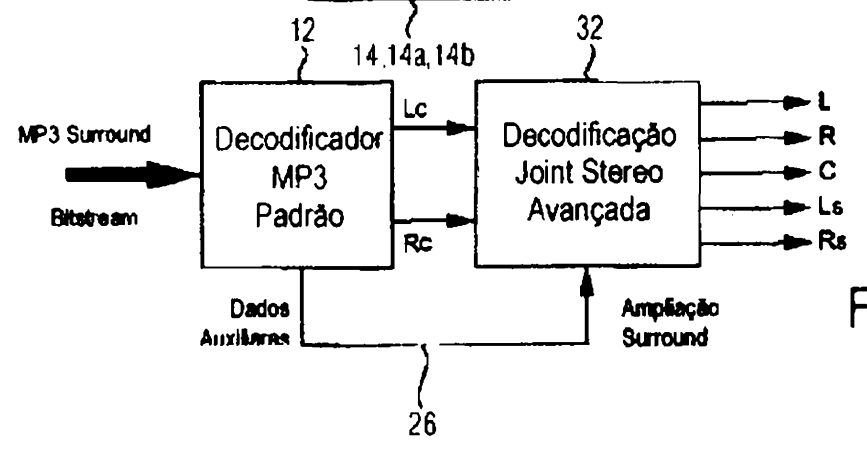
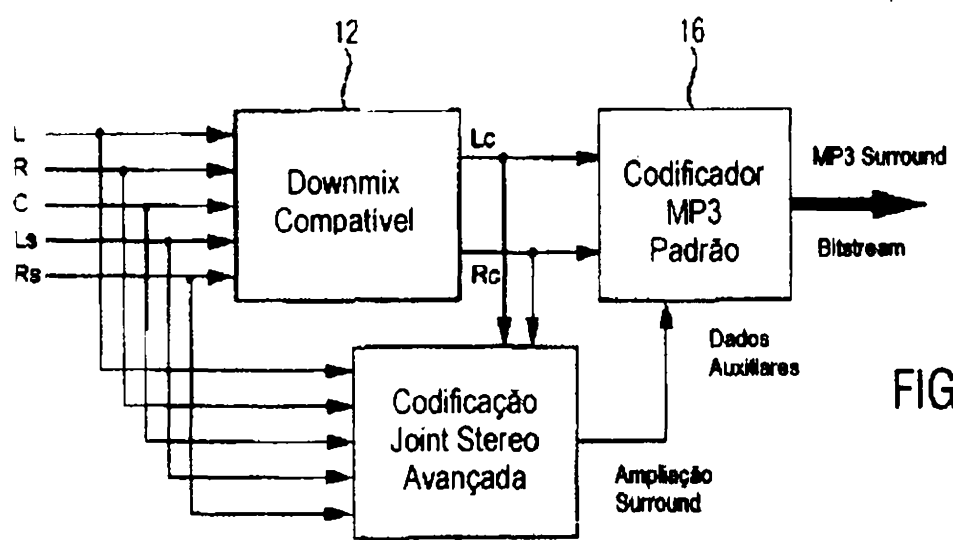
todos os sinais são sinais
no domínio do tempo

$a, b, t < 1$
preferido: $a = b = 0.7$
ou: $a = 0.5; b = 0.7$

FIG 7

informações
colaterais de
canal

	CANAL B	CANAL A
	informações colaterais/parâmetros para:	são determinadas usando:
l_i	L	L_C
l_{si}	L_S	L_C
r_i	R	R_C
r_{si}	R_S	R_C
C_i	C	$L_C + R_C$



7/7

FIG 10

