



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

(21) **PI0621499-1 A2**



(22) Data de Depósito: 18/05/2006
(43) Data da Publicação: 13/12/2011
(RPI 2136)

(51) *Int.Cl.:*
H04S 1/00
H04S 3/00

(54) **Título:** MÉTODO MELHORADO PARA FORMATAÇÃO DE SINAL EM RECONSTRUÇÃO DE ÁUDIO DE CANAIS MÚLTIPLOS

(30) **Prioridade Unionista:** 28/03/2006 US 60/787.096

(73) **Titular(es):** FRAUNHOFER-GESELLSCHAFT ZUR FORDERUNG DER ANGEWADTEN FORSCHUNG E.V.

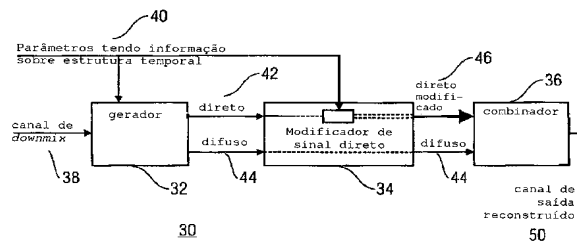
(72) **Inventor(es):** DISCH, SASCHA, HERRE, JURGEN, LINZMEIER, KARSTEN, POPP, HARALD

(74) **Procurador(es):** CRUZEIRO NEWMARC PATENTES E MARCAS LTDA

(86) **Pedido Internacional:** PCT EP2006004732 de 18/05/2006

(87) **Publicação Internacional:** WO 2007/110101de 04/10/2007

(57) **Resumo:** MÉTODO MELHORADO PARA FORMATAÇÃO DE SINAL EM RECONSTRUÇÃO DE ÁUDIO DE CANAIS MÚLTIPLOS. Canal de saída reconstruído, reconstruído com um reconstrutor de canais múltiplos usando pelo menos um canal de downmix derivado pela execução de downmix de uma pluralidade de canais originais, e usando uma representação de parâmetro incluindo informação adicional sobre uma estrutura fina temporal de um canal original, pode ser gerado usando um gerador (32) para gerar um componente de sinal direto (42), e um componente de sinal difuso (44) baseado no canal de downmix (38) é usado. Apenas o componente de sinal direto (42) é modificado (34), de modo que a estrutura fina temporal (40) do canal de saída reconstruído é adequada a uma estrutura final temporal desejada, indicada pela informação adicional sobre a estrutura fina temporal transmitida.





PI0621499-1

"MÉTODO MELHORADO PARA FORMATAÇÃO DE SINAL EM
RECONSTRUÇÃO DE ÁUDIO DE CANAIS MÚLTIPLOS"

Descrição

Campo da Invenção

5 A presente invenção se refere a um conceito de formatação de sinal melhorada em reconstrução de áudio de canais múltiplos e, especificamente, a uma nova abordagem de formatação de envelope.

Histórico da Invenção e Técnica Anterior

10 Desenvolvimento recente na codificação de áudio permite a recriação de uma representação de canais múltiplos de um sinal de áudio, baseada em um sinal estéreo (ou mono) e dados de controle correspondentes. Estes métodos diferem substancialmente de soluções baseadas em matriz mais antiga, tal como Dolby
15 Prologic, visto que dados de controle adicionais são transmitidos para controlar a recriação, também referidos como *upmix*, dos canais "surround" baseados nos canais mono ou estéreo transmitidos. Estes decodificadores de áudio de canais múltiplos paramétricos reconstroem N canais baseados em M canais
20 transmitidos, onde $N > M$, e nos dados de controle adicionais. O uso dos dados de controle adicionais resulta em uma taxa de dados significativamente menor que a transmissão de todos os N canais, tornando a codificação muito eficiente, enquanto, simultaneamente, é garantida compatibilidade com ambos, os dispositivos de M canais
25 e os dispositivos de N canais. Os M canais podem, também, ser um canal mono único, um canal estéreo, ou uma representação de canal 5.1. Dessa maneira, é possível ter um sinal original de canal 7.2, no qual foi efetuado *downmix* para um sinal compatível

retrogradamente com canal 5.1, e parâmetros de áudio espacial que permitem que um decodificador de áudio espacial reproduza uma versão bastante similar dos canais 7.2 originais, em uma pequena elevação de taxa de bit adicional.

5 Estes métodos de codificação de surround paramétrico usualmente compreendem uma parametrização do sinal surround baseada em parâmetros de variante de tempo e frequência ILD (Inter Channel Level Difference [Diferença de Nível Entre os Canais]) e ICC (Inter Channel Coherence [Coerência Entre Canais]).

10 Estes parâmetros descrevem, por exemplo, proporções de energia e correlações entre pares de canais do sinal de canais múltiplos originais. No processo de decodificação, o sinal de canais múltiplos recriado é obtido pela distribuição da energia dos canais de *downmix* recebidos entre todos os pares de canais

15 conforme descrito pelos parâmetros ILD transmitidos. Entretanto, visto que um sinal de canais múltiplos pode ter distribuição de energia igual entre todos os canais embora os sinais nos diferentes canais sejam bastante diferentes, dessa maneira provendo uma impressão de audição de um som muito amplo, a

20 amplitude correta é obtida pela mixagem dos sinais com versões descorrelacionadas dos mesmos, conforme descrito pelo parâmetro ICC.

A versão descorrelacionada do sinal, frequentemente referida também como sinal molhado ou difuso, é

25 obtida através da passagem do sinal através de um reverberador, tal como um filtro de passagem total. Uma forma simples de descorrelação é a aplicação de um atraso específico ao sinal. De forma geral, existe uma grande quantidade de reverberadores

diferentes conhecidos na técnica, a implementação precisa do reverberador usado é de mínima importância.

A saída do descorrelacionador tem um tempo de resposta que é usualmente muito plano. Assim, um sinal de entrada Dirac provê um arrebatamento de ruído degradante. Ao mixar o sinal descorrelacionado e o sinal original, isto ocorre para alguns tipos de sinais transientes como sinais de aplauso, é importante executar algum processamento posterior no sinal para evitar a capacidade de percepção de artefatos introduzidos adicionalmente que podem resultar em um tamanho de sala percebido maior e em um tipo de artefatos pré-eco.

De forma geral, a invenção se refere a um sistema que representa áudio de canais múltiplos como uma combinação de dados de *downmix* de áudio (por exemplo, um ou dois canais) e dados de canais múltiplos paramétricos relacionados. Neste esquema (por exemplo, em uma codificação de sinal de som binauricular) uma corrente de dados de *downmix* de áudio é transmitida, onde pode ser observado que a forma mais simples de *downmix* é simplesmente a adição dos sinais diferentes de um sinal de canais múltiplos. Este sinal (sinal de soma) é acompanhado por uma corrente de dados de canais múltiplos paramétricos (informação de lado). A informação de lado compreende, por exemplo, um ou mais dos tipos de parâmetros discutidos acima para descrever a inter-relação espacial dos canais originais do sinal de canais múltiplos. Em um sentido, o esquema de canais múltiplos paramétricos age como um pré-/pós-processador para extremidade de envio/recepção dos dados de *downmix*, por exemplo, tendo o sinal de soma e a informação de lado. Deve ser observado que o sinal de soma dos dados de *downmix*

pode, adicionalmente, ser codificado usando qualquer codificador de áudio ou voz.

Visto que a transmissão de sinais de canais múltiplos sobre transportadores com banda larga baixa foi se tornando mais e mais popular, estes sistemas, também conhecidos como "codificação de áudio espacial", "MPEG surround", foram se tornando mais bem desenvolvidos.

As publicações a seguir são conhecidas no contexto destas tecnologias:

10 [1] C. Faller e F. Baumgarte, "Efficient representation of spatial audio using perceptual parametrization," no Proc. IEEE WASPAA, Mohonk, NY, Outubro de 2001.

[2] F. Baumgarte e C. Faller, "Estimation of auditory spatial cues for binaural cue coding," no Proc. ICASSP 2002, Orlando, FL, Maio de 2002.

[3] C. Faller e F. Baumgarte, "Binaural cue coding: a novel and efficient representation of spatial audio," no Proc. ICASSP 2002, Orlando, FL, Maio de 2002.

20 [4] F. Baumgarte e C. Faller, "Why binaural cue coding is better than intensity stereo coding," no Proc. AES 112th Conv., Munique, Alemanha, Maio de 2002.

[5] C. Faller e F. Baumgarte, "Binaural cue coding applied to stereo and multi-channel audio compression," no Proc. AES 112th Conv., Munique, Alemanha, Maio de 2002.

25 [6] F. Baumgarte e C. Faller, "Design and evaluation of binaural cue coding," no AES 113th Conv., Los Angeles, CA, Outubro de 2002.

[7] C. Faller e F. Baumgarte, "Binaural cue

coding applied to audio compression with flexible rendering," no Proc. AES 113th Conv., Los Angeles, CA, Outubro de 2002.

[8] J. Breebaart, J. Herre, C. Faller, J. Rödén, F. Myburg, S. Disch, H. Purnhagen, G. Hoto, M. Neusinger, K. Kjörling, W. Oomen: "MPEG Spatial Audio Coding / MPEG Surround: Overview and Current Status", 119th AES Convention, Nova York 2005, Pré-impressão 6599

[9] J. Herre, H. Purnhagen, J. Breebaart, C. Faller, S. Disch, K. Kjörling, E. Schuijers, J. Hilpert, F. Myburg, "The Reference Model Architecture for MPEG Spatial Audio Coding", 118th AES Convention, Barcelona 2005, Pré-impressão 6477

[10] J. Herre, C. Faller, S. Disch, C. Ertel, J. Hilpert, A. Hoelzer, K. Linzmeier, C. Spenger, P. Kroon: "Spatial Audio Coding: Next-Generation Efficient and Compatible Coding of Multi-Channel Audio", 117th AES Convention, São Francisco 2004, Pré-impressão 6186

[11] J. Herre, C. Faller, C. Ertel, J. Hilpert, A. Hoelzer, C. Spenger: "MP3 Surround: Efficient and Compatible Coding of Multi-Channel Audio", 116th AES Convention, Berlin 2004, Pré-impressão 6049.

Uma técnica relacionada, focada na transmissão de dois canais por meio de um sinal mono transmitido é denominada de "estéreo paramétrico" e, por exemplo, descrita mais extensamente nas publicações a seguir:

[12] J. Breebaart, S. van de Par, A. Kohlrausch, E. Schuijers, "High-Quality Parametric Spatial Audio Coding at Low Bitrates", AES 116th Convention, Berlin, Pré-impressão 6072, Maio de 2004

[13] E. Schuijers, J. Breebaart, H. Purnhagen, J. Engdegard, "Low Complexity Parametric Stereo Coding", AES 116th Convention, Berlin, Pré-impressão 6073, Maio de 2004.

Em um decodificador de áudio espacial, o *upmix* de 5 canais múltiplos é computado de uma parte do sinal direto e de uma parte de sinal difuso, que é derivada por meio de uma descorrelação da parte direta, conforme já foi mencionado acima. Dessa maneira, no geral, a parte difusa tem um envelope temporal diferente da parte direta. O termo "envelope temporal" descreve 10 neste contexto a variação da energia ou amplitude do sinal em relação ao tempo. O envelope temporal diferente conduz a artefatos (pré- e pós-ecos, mancha ("smearing") temporal) nos sinais de *upmix* para sinais de entrada que têm uma imagem estéreo ampla e, simultaneamente, uma estrutura de envelope transiente. Sinais 15 transientes geralmente são sinais que variam grandemente em um curto período de tempo.

Provavelmente, os exemplos mais importantes para esta classe de sinais são os sinais tipo aplauso, que estão frequentemente presentes em gravações ao vivo.

20 De modo a evitar artefatos causados pela introdução de som difuso/descorrelacionado com um envelope temporal inapropriado no sinal de *upmix*, uma série de técnicas foi proposta:

O pedido Norte Americano 11/006.492 ("Diffuse 25 Sound Shaping for BCC Schemes and The Like") mostra que a qualidade de percepção de sinais transientes críticos pode ser melhorada pela formatação do envelope temporal do sinal difuso para corresponder com o envelope temporal do sinal direto.

Esta abordagem já havia sido introduzida na tecnologia MPEG surround por ferramentas diferentes, tais como "formatação de envelope temporal" (TES) e "processamento temporal" (TP). Visto que o envelope temporal alvo do sinal difuso é derivado do envelope do sinal de *downmix* transmitido, este método não requer que informação de lado adicional seja transmitida. Entretanto, como uma consequência, a estrutura fina temporal do som difuso é igual para todos os canais de saída. Visto que a parte de sinal direto, que é diretamente derivada do sinal de *downmix* transmitido, também tem um envelope temporal similar, este método pode melhorar a qualidade de percepção de sinais do tipo aplauso em termos de acentuação ("crispness"). Entretanto, como o sinal direto e o sinal difuso têm envelopes temporais diretos para todos os canais, estas técnicas podem aumentar a qualidade subjetiva de sinais do tipo aplauso, mas não podem melhorar a distribuição espacial de eventos de aplauso únicos no sinal, visto que isto só seria possível quando um canal reconstruído fosse muito mais intenso na ocorrência do sinal transiente do que os outros canais, o que é impossível tendo sinais que partilham basicamente o mesmo envelope temporal.

Um método alternativo para superar o problema é descrito no pedido Norte Americano 11/006.482 ("Individual Channel Shaping for BCC Schemes and The Like"). Esta abordagem emprega informação de lado de banda ampla temporal de grão fino que é transmitida pelo codificador para executar uma formatação temporal fina de ambos os sinais, o sinal direto e o sinal difuso. Evidentemente, esta abordagem permite uma estrutura fina temporal que é individual para cada canal de saída e, dessa maneira, é

capaz de acomodar também sinais para os quais eventos transientes ocorrem apenas em um subconjunto dos canais de saída. Uma variação adicional desta abordagem é descrita no pedido Norte Americano 60/726.389 ("Methods for Improved Temporal and Spatial Shaping of
5 Multi-Channel Audio Signals"). Ambas as abordagens discutidas para aumentar a qualidade de percepção dos sinais codificados transientes compreendem uma formatação temporal do envelope do sinal difuso que objetiva corresponder com um envelope temporal de sinais diretos correspondente.

10 Embora ambos os métodos da técnica anterior descritos anteriormente possam aumentar a qualidade subjetiva de sinais tipo aplauso em termos de acentuação, apenas a última abordagem pode melhorar também a redistribuição espacial do sinal reconstruído. Ainda, a qualidade subjetiva dos sinais de aplauso
15 sintetizados permanece insatisfatória, visto que a formatação temporal de ambas as combinações de sons secos e difusos conduz a distorções características (os ataques das palmas individuais são percebidos como "frouxos" quando apenas uma formatação temporal solta é executada, ou distorções são introduzidas se formatação
20 com uma resolução temporal muito elevada for aplicada ao sinal). Isto se torna evidente quando um sinal difuso é simplesmente uma cópia retardada do sinal direto. Então, o sinal difuso misturado ao sinal direto provavelmente terá uma composição espectral diferente que aquela do sinal direto. Dessa maneira, mesmo se o
25 envelope for escalonado para corresponder ao envelope do sinal direto, contribuições espectrais diferentes, não originárias diretamente do sinal original, estarão presentes no sinal reconstruído. As distorções introduzidas podem se tornar ainda

piores quando a parte do sinal difuso é enfatizada (com volume aumentado) durante a reconstrução, quando o sinal difuso é escalonado para corresponder ao envelope do sinal direto.

Sumário da Invenção

5 O objetivo da presente invenção é prover um conceito de formatação de sinal melhorada em reconstrução de canais múltiplos.

Este objetivo é atingido por um aparato de acordo com as reivindicações 1 ou 29, um método de acordo com a
10 reivindicação 28 e um programa de computador de acordo com a reivindicação 30.

A presente invenção é baseada na descoberta de que um canal de saída reconstruído, reconstruído com um reconstrutor de canais múltiplos usando pelo menos um canal de
15 *downmix* derivado por *downmix* de uma pluralidade de canais originais e usando uma representação de parâmetro que inclui informação adicional em uma estrutura (fina) temporal de um canal original, pode ser reconstruído eficientemente com alta qualidade quando um gerador para gerar um componente de sinal direto e um
20 componente de sinal difuso baseado no canal de *downmix* é usado. A qualidade pode ser essencialmente melhorada se apenas o componente de sinal direto for modificado, de modo que a estrutura fina temporal do canal de saída reconstruído esteja adequada a uma estrutura fina temporal desejada, indicada pela informação
25 adicional na estrutura fina temporal transmitida.

Em outras palavras, escalonar as partes de sinal direto derivadas diretamente do sinal de *downmix*, dificilmente introduz artefatos adicionais no momento que um sinal transiente

ocorre. Quando, como na técnica anterior, a parte de sinal molhado é escalonada para corresponder a um envelope desejado, pode, muito bem, ser a ocasião em que o sinal transiente original no canal reconstruído é mascarado por um sinal difuso enfatizado misturado ao sinal direto, o que será descrito abaixo mais extensivamente.

A presente invenção supera este problema por meio apenas do escalonamento do componente de sinal direto, dessa maneira eliminando a oportunidade de introduzir artefatos adicionais ao custo de transmissão de parâmetros adicionais para descrever o envelope temporal dentro da informação de lado.

De acordo com uma configuração da presente invenção, parâmetros de escalonamento de envelope são derivados usando uma representação do sinal direto e do sinal difuso com um espectro clareado ("whitened spectrum"), isto é, onde partes espectrais diferentes do sinal possuem energias quase idênticas. As vantagens do uso de espectros clareados são duplicadas. Uma, por um lado, usando um espectro clareado como uma base para o cálculo de um fator de escalonamento usado para escalonar o sinal direto, permite a transmissão de apenas um parâmetro por fenda de tempo incluindo informação sobre a estrutura temporal. Como é usual em codificação de áudio de canais múltiplos que sinais sejam processados dentro de numerosas bandas de frequência, esta característica ajuda a diminuir o número de informação de lado adicionalmente necessário e, dessa maneira, o aumento de taxa de bit para a transmissão do parâmetro adicional. Tipicamente, outros parâmetros tais como ICLD e ICC são transmitidos uma vez por quadro de tempo e banda de parâmetro. Visto que o número de bandas de parâmetro pode ser maior que 20, é uma vantagem importante ter

que transmitir apenas um único parâmetro por canal. De forma geral, em codificação de canais múltiplos, sinais são processados em uma estrutura de quadro, isto é, em entidades tendo vários valores de amostragem, por exemplo, 1024 por quadro. Além do mais, conforme já foi mencionado, os sinais são divididos em várias porções espectrais antes de serem processados, de modo que, finalmente, tipicamente apenas um parâmetro ICC e ICLD é transmitido por quadro e porção espectral do sinal.

A segunda vantagem de usar apenas um parâmetro é fisicamente motivada, visto que os sinais transientes em questão naturalmente possuem espectros amplos. Portanto, para contabilizar a energia dos sinais transientes dentro dos canais únicos corretamente, é muito apropriado usar espectros clareados para o cálculo de fatores de escalonamento de energia.

Em uma configuração adicional da presente invenção, o conceito inventivo de modificação do componente de sinal direto é aplicado apenas para uma porção espectral do sinal acima de certo limite espectral na presença de sinais residuais adicionais. Isto se deve ao fato dos sinais residuais juntamente com o sinal de *downmix* permitirem uma reprodução com qualidade superior dos canais originais.

Sumarizando, o conceito inventivo é projetado para prover qualidade espacial e temporal aumentada com relação às abordagens da técnica anterior, evitando os problemas associados com aquelas técnicas. Portanto, informação de lado é transmitida para descrever a estrutura de envelope de tempo fina dos canais individuais e, dessa maneira, permite formatação temporal/espacial fina dos sinais de canal de *upmix* no lado do decodificador. O

método inventivo descrito neste documento é baseado nas descobertas/considerações a seguir:

• Sinais do tipo aplauso podem ser vistos como compostos de palmas próximas únicas, distintas, e um ambiente do tipo ruído originário de palmas distantes muito densas.

• Em um decodificador de áudio espacial, a melhor aproximação das palmas próximas em termos de envelope temporal é o sinal direto. Portanto, apenas o sinal direto é processado pelo método inventivo.

10 • Visto que o sinal difuso representa principalmente a parte do ambiente do sinal, qualquer processamento em uma resolução temporal fina provavelmente introduzirá artefatos de distorção e modulação (mesmo se certo aumento subjetivo de "acentuação" de aplauso puder ser obtido por
15 esta técnica). Como uma consequência destas considerações, o sinal difuso é intocado (isto é, não está sujeito à formatação de tempo fina) pelo processamento inventivo.

• No entanto, o sinal difuso contribui para o equilíbrio da energia do sinal de *upmix*. O método inventivo
20 contabiliza isto pelo cálculo de um fator de escalonamento de transmissão da informação transmitida que deve ser aplicado unicamente à parte de sinal direto. Este fator modificado é escolhido de modo que a energia geral em um determinado intervalo de tempo seja igual dentro de certos limites, como se o fator
25 original tivesse sido aplicado a ambas, a parte direta e parte difusa do sinal neste intervalo.

• Usando o método inventivo, melhor qualidade de áudio subjetiva é obtida se a resolução espectral dos sinais de

som espaciais for escolhida para ser baixa - por exemplo, "largura de banda total" - para garantir a preservação da integridade espectral dos transientes contidos no sinal. Neste caso, o método proposto não aumenta necessariamente a taxa de bit média de
 5 informação de lado espacial, visto que a resolução espectral é seguramente trocada por resolução temporal.

O melhoramento na qualidade subjetiva é atingido pela amplificação ou amortecimento ("formatação") da parte seca do sinal em relação apenas ao tempo e, dessa maneira:

10 • Aumentar a qualidade transiente pelo fortalecimento da parte de sinal direto no local transiente, enquanto distorção adicional é evitada, originária de um sinal difuso com envelope temporal não apropriado.

15 • Melhorar a localização espacial pela ênfase da parte direta w.r.t. da parte difusa na origem espacial de um evento transiente e amortecimento em relação à parte difusa nas posições de colocação de uma fonte de som em um campo estéreo ("panning") distantes.

Breve Descrição dos Desenhos

20 A Figura 1 mostra um diagrama de bloco de um codificador de canais múltiplos e um decodificador correspondente;

A Figura 1b mostra um desenho esquematizado de uma reconstrução de sinal usando sinais descorrelacionados;

25 A Figura 2 mostra um exemplo de um reconstrutor de canais múltiplos inventivo;

A Figura 3 mostra um exemplo adicional de um reconstrutor de canal múltiplo inventivo;

A Figura 4 mostra um exemplo de representações de

banda de parâmetro usadas para identificar bandas de parâmetro diferentes dentro de um esquema de decodificação de canais múltiplos;

A Figura 5 mostra um exemplo de um decodificador de canais múltiplos inventivo; e

A Figura 6 mostra um diagrama de bloco detalhando um exemplo de um método inventivo de reconstrução de um canal de saída.

Descrição Detalhada das Configurações Adicionais

A Figura 1 mostra um exemplo de codificação de dados de áudio de canais múltiplos de acordo com a técnica anterior, para ilustrar de forma mais clara o problema solucionado pelo conceito inventivo.

De forma geral, em um lado do codificador, um sinal de canais múltiplos original é inserido no codificador de canais múltiplos, derivando informação de lado indicando a distribuição espacial dos vários canais dos sinais de canais múltiplos originais com relação uns aos outros. Além da geração da informação de lado, um codificador de canais múltiplos gera um ou mais sinais de soma, que sofreram *downmix* a partir do sinal de canais múltiplos original.

Configurações famosas amplamente usadas são denominadas de configurações 5-1-5 e 5-2-5. Na configuração 5-1-5 o codificador gera um único sinal de soma monofônico a partir de cinco canais de entrada e, assim, um decodificador correspondente tem que gerar cinco canais reconstruídos de um sinal de canais múltiplos reconstruído. Na configuração 5-2-5, o codificador gera dois canais de *downmix* de cinco canais de

entrada, o primeiro canal dos canais de *downmix* tipicamente retendo informação sobre um lado esquerdo ou um lado direito, e o segundo canal dos canais de *downmix* retendo informação do outro lado.

5 Parâmetros de amostra que descrevem a distribuição espacial dos canais originais são, conforme indicado, por exemplo, na Figura 1, os parâmetros previamente introduzidos ICLD e ICC.

10 Pode ser observado que dentro da análise que deriva a informação de lado 14, as amostras dos canais originais do sinal de canais múltiplos 10 são tipicamente processadas em domínios de sub-bandas representando um intervalo de frequência específico dos canais originais. Um intervalo de frequência único é indicado por k . Em algumas aplicações, os canais de entrada
15 podem ser filtrados por um banco de filtro híbrido antes do processamento, isto é, as bandas de parâmetro k podem ser adicionalmente subdivididas, cada subdivisão denotada por k .

 Além do mais, o processamento dos valores da amostra que descrevem um canal original, é feito de uma maneira do
20 tipo "quadro" dentro de cada banda de parâmetro única, isto é, várias amostras consecutivas formam um quadro de duração finita. Os parâmetros BCC mencionados acima tipicamente descrevem um quadro total.

 Um parâmetro relacionado de alguma maneira com a
25 presente invenção, e já conhecido na técnica, é o parâmetro ICLD que descreve a energia contida dentro de um quadro de sinal de um canal com relação aos quadros correspondentes de outros canais dos canais múltiplos ou sinais originais.

Comumente, a geração de canais adicionais para derivar uma reconstrução de um sinal de canais múltiplos de um sinal de soma transmitido apenas é atingida com a ajuda de sinais descorrelacionados, sendo derivados do sinal de soma usando descorrelacionadores ou reverberadores. Para uma aplicação típica, a frequência de amostra discreta pode ser de 44,100 kHz, de modo que uma amostra única representa um intervalo de comprimento finito de aproximadamente 0,02 ms de um canal original. Pode ser observado que, usando bancos de filtro, o sinal é dividido em numerosas partes de sinal, cada uma representando um intervalo de frequência finito do sinal original. Para compensar um possível aumento nos parâmetros que descrevem o canal, a resolução de tempo é normalmente diminuída, de modo que uma porção de tempo de comprimento finito descrita por uma amostra única dentro de um domínio de banco de filtro pode aumentar para mais que 0,5 ms. O comprimento típico de um quadro pode variar entre 10 e 15 ms.

A derivação do sinal descorrelacionado pode fazer uso de estruturas de filtro diferentes e/ou atrasos ou combinações dos mesmos sem limitar o escopo da invenção. Pode ser, além do mais, observado que o espectro inteiro não tem, necessariamente, que ser usado para derivar os sinais descorrelacionados. Por exemplo, apenas porções espectrais acima de um limite inferior espectral (valor específico de k) do sinal de soma (sinal de *downmix*) podem ser usadas para derivar os sinais descorrelacionados usando atrasos e/ou filtros. Um sinal descorrelacionado, dessa maneira, geralmente descreve um sinal derivado do sinal de *downmix* (canal de *downmix*) de modo que um coeficiente de correlação, quando derivado usando o sinal

descorrelacionado e o canal de *downmix*, significativamente desvia da unidade, por exemplo, em 0,2.

A Figura 1b provê um exemplo extremamente simplificado do *downmix* e do processo de reconstrução durante codificação de áudio de canais múltiplos para explicar o grande benefício do conceito inventivo de escalonamento apenas do componente do sinal direto durante a reconstrução de um canal de um sinal de canais múltiplos. Para a descrição a seguir, algumas simplificações são pressupostas. A primeira simplificação é que o *downmix* de um canal esquerdo e um canal direito é uma adição simples das amplitudes dentro dos canais. A segunda simplificação potente é que a correlação é pressuposta ser um atraso simples do sinal total.

De acordo com estas pressuposições, um quadro de um canal esquerdo 21a e um canal direito 21b deve ser codificado. Conforme indicado no eixo x das janelas mostradas, em codificação de áudio de canais múltiplos, o processamento é tipicamente executado em valores de amostra, amostrados com uma frequência de amostra fixa. Isto será, para facilidade de explanação, desconsiderado adicionalmente no breve sumário a seguir.

Conforme já mencionado, no lado do codificador, um canal esquerdo e direito é combinado (*downmix*) em um canal de *downmix* 22 que é para ser transmitido para o decodificador. No lado do decodificador, um sinal descorrelacionado 23 é derivado do canal de *downmix* transmitido 22, que é a soma do canal esquerdo 21a e do canal direito 21b neste exemplo. Conforme já foi explicado, a reconstrução do canal esquerdo é, então, executada a partir de quadros de sinal derivados do canal de *downmix* 22 e do

sinal descorrelacionado 23.

Pode ser observado que cada quadro único está sendo submetido a escalonamento global antes da combinação, conforme indicado pelo parâmetro ICLD, que relaciona as energias dentro dos quadros individuais de canais únicos com a energia dos quadros correspondentes dos outros canais de um sinal de canais múltiplos.

Como é pressuposto no presente exemplo, que energias iguais estão contidas dentro do quadro do canal esquerdo 21a e do quadro do canal direito 21b, o canal de *downmix* transmitido 22 e o sinal descorrelacionado 23 são escalonados por um fator de aproximadamente 0,5 antes da combinação. Isto é, quando o *upmix* é igualmente simples ao *downmix*, isto é, somando os dois sinais, a reconstrução do canal esquerdo original 21a é a soma do canal de *downmix* escalonado 24a e do sinal descorrelacionado escalonado 24b.

Devido à soma para transmissão e ao escalonamento devido ao parâmetro ICLD, o sinal para a taxa de fundo do sinal transiente deveria ser diminuído por um fator de aproximadamente 2. Além do mais, ao adicionar simplesmente os dois sinais, um tipo de eco adicional de artefato seria introduzido na posição da estrutura transiente atrasada no sinal descorrelacionado escalonado 24b.

Conforme indicado na Figura 1b, a técnica anterior supera o problema de eco pelo escalonamento da amplitude do sinal descorrelacionado escalonado 24b para fazer que ele corresponda ao envelope do canal transmitido escalonado 24a, conforme indicado pelas linhas pontilhadas no quadro 24b. Devido

ao escalonamento, a amplitude na posição do sinal transiente original no canal esquerdo 21a pode ser aumentada. Entretanto, a composição espectral do sinal descorrelacionado na posição de escalonamento no quadro 24b é diferente da composição espectral do sinal transiente original. Portanto, artefatos audíveis são introduzidos no sinal, mesmo se a intensidade geral do sinal puder ser bem reproduzida.

A grande vantagem da presente invenção é que a presente invenção não escala apenas um componente de sinal direto de reconstrução. Visto que este canal não tem um componente de sinal correspondente ao sinal transiente original tendo a composição espectral correta e o escalonamento de tempo correto, escalonar apenas o canal de *downmix* renderá um sinal reconstruído que reconstrói o evento transiente original com alta precisão. Este é o caso visto que apenas partes do sinal são enfatizadas pelo escalonamento que tem a mesma composição espectral do sinal transiente original.

A Figura 2 mostra um diagrama de bloco de um exemplo de um reconstrutor de canais múltiplos da invenção, para detalhar a característica principal do conceito inventivo.

A Figura 2 mostra um reconstrutor de canais múltiplos 30, tendo um gerador 32, um modificador e um combinador de sinal direto 36. O gerador 32 recebe um canal de *downmix* 38 em que foi efetuado *downmix* de uma pluralidade de canais originais e uma representação de parâmetro 40 que inclui informação sobre uma estrutura temporal de um canal original.

O gerador gera um componente de sinal direto 42 e um componente de sinal difuso 44 baseados no canal de *downmix*.

O modificador de sinal direto 34 recebe, também, o componente de sinal direto 42, como o componente de sinal difuso 44, e adicionalmente, a representação de parâmetro 40 tendo a informação sobre uma estrutura temporal do canal original. De acordo com a presente invenção, o modificador de sinal direto 34 modifica apenas o componente de sinal direto 42 usando a representação de parâmetro para derivar um componente de sinal direto modificado 46.

O componente de sinal direto modificado 46 e o componente de sinal difuso 44, que não é alterado pelo modificador de sinal direto 34, são inseridos no combinador 36 que combina o componente de sinal direto modificado 46 e o componente de sinal difuso 44 para obter um canal de saída reconstruído 50.

Apenas pela modificação do componente de sinal direto 42 derivado do canal de *downmix* transmitido 38 sem reverberação (descorrelação), é possível reconstruir o envelope de tempo para o canal de saída reconstruído correspondendo muito proximamente a um envelope de tempo do canal original subjacente sem introduzir artefatos adicionais e distorções audíveis, como nas tecnologias da técnica anterior.

Como será discutido mais detalhadamente na descrição da Figura 3, a formatação de envelope da invenção recupera o envelope de banda ampla do sinal de saída sintetizado. Isto compreende um procedimento de *upmix* modificado, seguido pelo achatamento do envelope e reformatação da porção do sinal direto de cada canal de saída. Para reformatação, informação de lado de envelope de banda ampla paramétrica contida na corrente de bit da representação de parâmetro é usada. Esta informação de lado

consiste, de acordo com uma configuração da presente invenção, de proporções (EnvRatio) relacionando o envelope de sinal de *downmix* transmitido ao envelope do sinal do canal de entrada original. No decodificador, fatores de ganho são derivados destas proporções para serem aplicados ao sinal direto em cada fenda de tempo em um quadro de um determinado canal de saída. A porção de som difusa de cada canal não é alterada de acordo como conceito inventivo.

A configuração preferida da presente invenção mostrada no diagrama de bloco da Figura 3 é um reconstrutor de canais múltiplos 60 modificado para se adequar ao fluxo de sinal do decodificador de um decodificador MPEG espacial.

O reconstrutor de canais múltiplos 60 compreende um gerador 62 para gerar um componente de sinal direto 64 e um componente de sinal difuso 66 usando um canal de *downmix* 68 derivado pelo *downmix* de uma pluralidade de canais originais e uma representação de parâmetro 70 tendo informação sobre propriedades espaciais de canais originais do sinal de canais múltiplos, conforme usado dentro da codificação MPEG. O reconstrutor de canais múltiplos 60 compreende ainda um modificador de sinal direto 68, que recebe o componente de sinal direto 64, o componente de sinal difuso 66, o sinal de *downmix* 69 e informação de lado de envelope adicional 72 como entrada.

O modificador de sinal direto provê em sua saída de modificador 73, o componente de sinal direto modificado, modificado conforme descrito em maiores detalhes abaixo.

O combinador 74 recebe o componente de sinal direto modificado e o componente de sinal difuso para obter o canal de saída reconstruído 76.

Conforme mostrado na Figura, a presente invenção pode ser facilmente implementada em ambientes de canais múltiplos já existentes. Aplicação geral do conceito inventivo dentro deste esquema de codificação poderia ser ativada e desativada de acordo com alguns parâmetros adicionalmente transmitidos dentro da corrente de bit de parâmetro. Por exemplo, um marcador adicional **bsTempShapeEnable** poderia ser introduzido, que indica, quando definido como 1, que o uso do conceito inventivo é requerido.

Além do mais, um marcador adicional poderia ser introduzido especificamente definindo a necessidade da aplicação do conceito inventivo em um canal por base de canal. Portanto, um marcador adicional pode ser usado, denominado, por exemplo, **bsEnvShapeChannel**. Este marcador, disponível para cada canal individual, pode, então, indicar o uso do conceito inventivo, quando definido como 1.

Pode ser observado ainda que para facilidade de apresentação, apenas uma configuração de dois canais é descrita na Figura 3. Obviamente, a presente invenção não objetiva ser limitada apenas às configurações de dois canais. Além do mais, qualquer configuração de canal pode ser usada em conexão com o conceito inventivo. Por exemplo, cinco ou sete canais de entrada podem ser usados em conexão com a formatação de envelope avançada inventiva.

Quando o conceito inventivo é aplicado dentro de um esquema de codificação MPEG, conforme indicado na Figura 3, e a aplicação do conceito inventivo é sinalizada pela definição de **bsTempShapeEnable** igual a 1, componentes de sinal direto e difuso são sintetizados separadamente pelo gerador 62 usando uma pós-

mixagem modificada no domínio de sub-banda híbrido de acordo com a fórmula a seguir:

$$\mathbf{y}_{direct}^{n,k} = \mathbf{M}^{n,k} \mathbf{w}_{direct}^{n,k} \quad 0 \leq k < K$$

$$\mathbf{y}_{diffuse}^{n,k} = \mathbf{M}^{n,k} \mathbf{w}_{diffuse}^{n,k} \quad 0 \leq k < K$$

5 Aqui e nos parágrafos a seguir, o vetor $\mathbf{w}_{m,k}$ descreve o vetor de n parâmetros de sub-banda híbridos para a sub-banda k do domínio de sub-banda. Conforme indicado pela equação acima, parâmetros de sinal direto e difuso \mathbf{y} são separadamente derivados no *upmix*. As saídas diretas retêm o componente de sinal
10 direto e o sinal residual, que é um sinal que pode estar adicionalmente presente na codificação MPEG. Saídas difusas provêm apenas o sinal difuso. De acordo com o conceito inventivo, apenas o componente de sinal direto é adicionalmente processado pela formatação de envelope guiada (a formatação de envelope
15 inventiva).

O processo de formatação de envelope emprega uma operação de extração de envelope em diferentes sinais. O processo de extração de envelope ocorrendo dentro do modificador de sinal direto 68 é descrita em maiores detalhes nos parágrafos a seguir,
20 visto que isto é uma etapa obrigatória antes da aplicação da modificação inventiva ao componente de sinal direto.

Conforme já foi mencionado, dentro do domínio de sub-banda híbrida, sub-bandas são denotadas como k . Várias sub-bandas k podem também estar organizadas nas bandas de parâmetro k .

25 A associação de sub-bandas em bandas de parâmetro subjacentes à configuração da presente invenção discutida abaixo, é provida na tabela da Figura 4.

Primeiramente, para cada fenda em um quadro, as

energias E_{slot}^κ de certas bandas de parâmetro κ são calculadas com $y^{n,k}$ sendo um sinal de entrada de sub-banda híbrida.

$$E_{slot}^\kappa(n) = \sum_k y^{n,\tilde{k}} (y^{n,\tilde{k}})^* \quad \tilde{k} = \{k | \bar{\kappa}(k) = \kappa\} \quad \forall \kappa_{start} < \kappa < \kappa_{stop}$$

com $\kappa_{start} = 10$ e $\kappa_{stop} = 18$

5 A soma inclui todo \tilde{k} sendo atribuído a uma banda de parâmetro κ de acordo com a Tabela A.1.

Subsequentemente, uma média de energia de longa duração \bar{E}_{slot}^κ para cada banda de parâmetro é calculada como

$$\bar{E}_{slot}^\kappa(n) = (1 - \alpha) E_{slot}^\kappa(n) + \alpha \bar{E}_{slot}^\kappa(n-1)$$

10
$$\alpha = \exp\left(-\frac{64}{0.4 \cdot 44100}\right)$$

Com α sendo um fator de peso correspondendo a uma passagem baixa IIR de primeira ordem (aproximadamente 400 ms de constante de tempo) e n denotando um índice de fenda de tempo. A energia média total suavizada (banda ampla) \bar{E}_{total} é calculada como

15 sendo

$$\bar{E}_{total}(n) = (1 - \alpha) E_{total}(n) + \alpha \bar{E}_{total}(n-1)$$

com

$$E_{total}(n) = \frac{1}{\kappa_{stop} - \kappa_{start} + 1} \sum_{\kappa=\kappa_{start}}^{\kappa_{stop}} E_{slot}^\kappa(n)$$

$$\alpha = \exp\left(-\frac{64}{0.4 \cdot 44100}\right)$$

20 Como pode ser visto das fórmulas acima, o envelope temporal é suavizado antes dos fatores de ganho serem derivados da representação suavizada dos canais. A suavização geralmente significa derivar uma representação suavizada de um canal original tendo gradientes diminuídos.

Como pode ser observado a partir das fórmulas acima, a operação de clareamento descrita é baseada em estimativas de energia total suavizada e estimativas de energia suavizada em sub-bandas, dessa maneira garantindo maior estabilidade das estimativas de envelope final.

A proporção destas energias é determinada para obter pesos para uma operação de clareamento espectral:

$$w^{\kappa}(n) = \frac{\bar{E}_{total}(n)}{\bar{E}_{slot}^{\kappa}(n) + \varepsilon}$$

A estimativa de envelope de banda ampla é obtida pela soma das contribuições pesadas das bandas de parâmetro, normalizando em uma média de energia de longa duração e cálculo da raiz quadrada

$$Env(n) = \sqrt{\frac{EnvAbs(n)}{\overline{Env}(n)}}$$

com

$$EnvAbs(n) = \sum_{\kappa=\kappa_{start}}^{\kappa_{stop}} w^{\kappa}(n) \cdot E_{slot}^{\kappa}(n)$$

$$\overline{Env}(n) = (1 - \beta) EnvAbs(n) + \beta \overline{Env}(n-1)$$

$$\beta = \exp\left(-\frac{64}{0.04 \cdot 44100}\right)$$

β é um fator de peso correspondendo a uma passagem baixa de primeira ordem IIR (aproximadamente 40 ms de constante de tempo).

Energia espectralmente clareada ou medições de amplitude são usadas como a base para o cálculo de fatores de escalonamento. Como pode ser visto das fórmulas acima, clareamento espectral significa alterar o espectro, de modo que a mesma energia ou amplitude média esteja contida dentro de cada banda

espectral da representação dos canais de áudio. Isto é mais vantajoso visto que os sinais transientes em questão têm espectros muito amplos de modo que é necessário usar informação total do espectro disponível integral para o cálculo de fatores de ganho, para não suprimir os sinais transientes com relação a outros sinais não transientes. Em outras palavras, sinais clareados espectralmente são sinais que possuem energia aproximadamente igual em bandas espectrais diferentes de sua representação espectral.

10 O modificador de sinal direto inventivo modifica o componente de sinal direto. Conforme já mencionado, processamento pode ser restrito a alguns índices de sub-banda iniciando com um índice de início, na presença de sinais residuais transmitidos. Além do mais, o processamento pode, de forma geral, ser restrito a índices de sub-bandas acima do índice de limite.

O processo de formatação de envelope consiste de um achatamento do envelope de som direto para cada canal de saída, seguido por uma reformatação na direção do envelope alvo. Isto resulta em uma curva de ganho sendo aplicada ao sinal direto de cada canal de saída se **bsEnvShapeChannel=1** for sinalizado para este canal na informação de lado.

O processamento é feito apenas para certas sub-bandas híbridas k :

$$k > 7$$

25 Na presença de sinais residuais transmitidos, k é escolhido para iniciar acima da maior banda residual envolvida no *upmix* do canal em questão.

Para configuração 5-1-5 o envelope alvo é obtido

pela estimativa do envelope do *downmix* transmitido Env_{Dmx} , conforme descrito na seção anterior, e subsequentemente escalonando o mesmo com as proporções $envRatio_{ch}$ de envelope transmitido e re-quantizado do codificador.

5 Então, uma curva de ganho $g_{ch}(n)$ para todas as fendas em um quadro é calculada para cada canal de saída pela estimativa de seu envelope Env_{ch} e relaciona o mesmo com o envelope alvo. Finalmente, esta curva de ganho é convertida em uma curva de ganho efetiva unicamente para escalonamento da parte direta do
10 canal de *upmix*:

$$ratio_{ch}(n) = \min(4, \max(0.25, g_{ch} + ampRatio_{ch}(n) \cdot (g_{ch} - 1)))$$

com

$$g_{ch}(n) = \frac{envRatio_{ch}(n) \cdot Env_{Dmx}(n)}{Env_{ch}(n)}$$

$$ampRatio_{ch}(n) = \frac{\sum_k |y_{ch,diffuse}^{n,k}|}{\sum_k |y_{ch,direct}^{n,k}| + \epsilon}$$

$$ch \in \{L, Ls, C, R, Rs\}$$

Para configuração 5-2-5, o envelope alvo para L e
15 Ls é derivado do envelope do sinal de *downmix* transmitido do canal esquerdo Env_{DmxL} , para R e Rs o envelope de *downmix* transmitido do canal direito é usado Env_{DmxR} . O canal central é derivado da soma dos envelopes do sinal de *downmix* transmitido esquerdo e direito.

A curva de ganho é calculada para cada canal de
20 saída pela estimativa de seu envelope $Env^{L,Ls,C,R,Rs}$ e relaciona o mesmo com o envelope alvo. Em uma segunda etapa esta curva de ganho é convertida em uma curva de ganho efetiva para escalonar unicamente a parte direta do canal de *upmix*:

$$ratio_{ch}(n) = \min(4, \max(0.25, g_{ch} + ampRatio_{ch}(n) \cdot (g_{ch} - 1)))$$

com

$$ampRatio_{ch}(n) = \frac{\sum_k |y_{ch,diffuse}^{n,k}|}{\sum_k |y_{ch,direct}^{n,k}| + \varepsilon}, \quad ch \in \{L, Ls, C, R, Rs\}$$

$$g_{ch}(n) = \frac{envRatio_{ch}(n) \cdot Env_{DmxL}(n)}{Env_{ch}(n)}, \quad ch \in \{L, Ls\}$$

$$g_{ch}(n) = \frac{envRatio_{ch}(n) \cdot Env_{DmxR}(n)}{Env_{ch}(n)}, \quad ch \in \{R, Rs\}$$

$$g_{ch}(n) = \frac{envRatio_{ch}(n) \cdot 0.5(Env_{DmxL}(n) + Env_{DmxR}(n))}{Env_{ch}(n)}, \quad ch \in \{C\}$$

Para todos os canais, a curva de ganho de ajuste
5 de envelope é aplicada se **bsEnvShapeChannel=1**.

$$\tilde{y}_{ch,direct}^k(n) = ratio_{ch}(n) \cdot y_{ch,direct}^k(n), \quad ch \in \{L, Ls, C, R, Rs\}$$

Ainda, o sinal direto é apenas copiado

$$\tilde{y}_{ch,direct}^k(n) = y_{ch,direct}^k(n), \quad ch \in \{L, Ls, C, R, Rs\}$$

Finalmente, o componente de sinal direto
10 modificado de cada canal individual tem que ser combinado com o
componente de sinal difuso do canal individual correspondente
dentro do domínio de sub-banda híbrida de acordo com a equação a
seguir:

$$y_{ch}^{n,k} = \tilde{y}_{ch,direct}^{n,k} + y_{ch,diffuse}^{n,k}, \quad ch \in \{L, Ls, C, R, Rs\}$$

15 Como pode ser visto a partir dos parágrafos
acima, o conceito inventivo ensina a melhorar a qualidade de
percepção e distribuição espacial de sinais tipo aplauso em um
decodificador de áudio espacial. O melhoramento é obtido pela
derivação de fatores de ganho com granularidade temporal fina de
20 escalonamento para escalonar apenas a parte direta do sinal de
upmix espacial. Estes fatores de ganho são derivados

essencialmente de informação de lado transmitidas e nível ou medições de energia do sinal direto e difuso no codificador.

Como o exemplo acima especificamente descreve o cálculo baseado em medições de amplitude, deveria ser observado
5 que o método inventivo não está restrito a isto, mas poderia também calcular, por exemplo, com medições de energia ou outras quantidades adequadas para descrever um envelope temporal de um sinal.

O exemplo acima descreve o cálculo para
10 configurações de canal 5-1-5 e 5-2-5. Naturalmente, o princípio descrito acima poderia ser aplicado analogamente, por exemplo, a configurações de canal 7-2-7 e 7-5-7.

A Figura 5 mostra um exemplo de um decodificador de áudio de canais múltiplos inventivo 100, que recebe um canal de
15 *downmix* 102 derivado pelo *downmix* de uma pluralidade de canais de um sinal de canais múltiplos original e uma representação de parâmetro 104 incluindo informação sobre uma estrutura temporal dos canais originais (frontal esquerdo, frontal direito, posterior esquerdo e posterior direito) do sinal de canais múltiplos
20 original. O decodificador de canais múltiplos 100 está tendo um gerador 106 para gerar um componente de sinal direto e um componente de sinal difuso para cada um dos canais originais subjacentes ao canal de *downmix* 102. O decodificador de canais múltiplos 100 compreende, adicionalmente, quatro modificadores de
25 sinal direto inventivos 108a a 108d para cada um dos canais a ser reconstruído, de modo que o decodificador de canais múltiplos envia quatro canais de saída (frontal esquerdo, frontal direito, posterior esquerdo e posterior direito) em suas saídas 112.

Embora o decodificador de canais múltiplos inventivo tenha sido detalhado usando uma configuração exemplificativa de quatro canais originais a serem reconstruídos, o conceito inventivo pode ser implementado em esquemas de áudio de 5 canais múltiplos tendo números arbitrários de canais.

A Figura 6 mostra um diagrama de bloco, detalhando o método inventivo de gerar um canal de saída reconstruído.

Em uma etapa de geração 110, um componente de 10 sinal direto e um componente de sinal difuso são derivados do canal de *downmix*. Em uma etapa de modificação 112, o componente de sinal direto é modificado usando parâmetros da representação de parâmetro tendo informação sobre uma estrutura temporal de um canal original.

15 Em uma etapa de combinação 114, o componente de sinal direto modificado e o componente de sinal difuso são combinados para obter um canal de saída reconstruído.

Dependendo de certos requisitos de implementação dos métodos inventivos, os métodos inventivos podem ser 20 implementados em hardware ou em software. A implementação pode ser executada usando um meio de armazenagem digital, em especial um disco, DVD ou um CD tendo sinais de controle legíveis eletronicamente armazenados nos mesmos, que cooperem com um sistema de computador programável de modo que os métodos 25 inventivos sejam executados. De forma geral, a presente invenção é, portanto, um produto de programa de computador com um código de programa armazenado em um dispositivo legível por máquina, o código de programa sendo operacional para executar os métodos

inventivos quando o produto de programa de computador é executado em um computador. Em outras palavras, os métodos inventivos são, portanto, um programa de computador tendo um código de programa para executar pelo menos um dos métodos inventivos quando o programa de computador estiver sendo executado em um computador.

Embora a descrição anterior tenha sido especificamente efetuada e mostrada com referência às configurações especificadas da mesma, será compreendido por aqueles especializados na técnica que várias outras modificações na forma e detalhes podem ser feitas sem se afastar do espírito e escopo da mesma. Deve ser compreendido que várias alterações podem ser feitas na adaptação às diferentes configurações sem se afastar dos conceitos amplos apresentados aqui e compreendidos pelas reivindicações a seguir.

REIVINDICAÇÕES

1. Reconstructor de canais múltiplos (30; 60) para gerar um canal de saída reconstruído (50; 76) usando pelo menos um canal de *downmix* (38; 68) derivado por *downmix* de uma pluralidade de canais originais e usando uma representação de parâmetro (40; 72), a representação de parâmetro (40; 72) incluindo informação sobre uma estrutura temporal de um canal original, caracterizado pelo fato de que compreende: um gerador (32; 62) para gerar um componente de sinal direto (42; 64) e um componente de sinal difuso (44; 66) para o canal de saída reconstruído (50; 76), baseado no canal de *downmix* (38; 68); um modificador de sinal direto (34; 69) para modificar o componente de sinal direto (42; 64) usando a representação de parâmetro (40; 72); e um combinador (36; 74) para combinar o componente de sinal direto modificado (46) e o componente de sinal difuso (44; 66) para obter o canal de saída reconstruído (50; 76).

2. Reconstructor de canais múltiplos, de acordo com a reivindicação 1, caracterizado pelo fato de que o gerador (32; 62) é operacional para gerar o componente de sinal direto (42; 64) usando apenas componentes do canal de *downmix* (38; 68).

3. Reconstructor de canais múltiplos (30; 60), de acordo com as reivindicações 1 ou 2, caracterizado pelo fato de que o gerador (32; 62) é operacional para gerar o componente de sinal difuso (44; 66) usando uma porção filtrada e/ou atrasada do canal de *downmix* (38; 68).

4. Reconstructor de canais múltiplos (30; 60), de acordo com qualquer uma das reivindicações de 1 a 3, caracterizado pelo fato de que o modificador de sinal direto (34; 69) é

operacional para usar informação sobre a estrutura temporal do canal original indicando a energia contida no canal original dentro de uma porção de tempo de comprimento finito do canal original.

5 5. Reconstructor de canais múltiplos (30; 60), de acordo com qualquer uma das reivindicações de 1 a 3, caracterizado pelo fato de que o modificador de sinal direto (34; 69) é operacional para usar informação sobre a estrutura temporal do canal original indicando uma amplitude média do canal original
10 dentro de uma porção de tempo de comprimento finito do canal original.

 6. Reconstructor de canais múltiplos (30; 60), de acordo com qualquer uma das reivindicações de 1 a 5, caracterizado pelo fato de que o combinador (36; 74) é operacional para
15 adicionar o componente de sinal direto modificado (46) e o componente de sinal difuso (44; 66) para obter o sinal reconstruído.

 7. Reconstructor de canais múltiplos, de acordo com qualquer uma das reivindicações de 1 a 6, no qual o
20 reconstructor de canais múltiplos é operacional para usar um primeiro canal de *downmix* tendo informação sobre um lado esquerdo da pluralidade de canais originais e um segundo canal de *downmix* (38; 68) tendo informação sobre um lado direito da pluralidade de canais originais, caracterizado pelo fato de que um primeiro canal
25 de saída reconstruído (50; 76) para um lado esquerdo é combinado usando apenas componentes de sinal direto e difuso gerados do primeiro canal de *downmix* e onde um segundo canal de saída reconstruído para um lado direito é combinado usando componentes

de sinal direto e difuso gerados apenas do segundo sinal de *downmix*.

8. Gerador de canais múltiplos (30; 60), de acordo com qualquer uma das reivindicações de 1 a 7, no qual o
5 modificador de sinal direto (34; 68) é operacional para modificar o sinal direto para porções de tempo de comprimento finito sendo mais curtas que as porções de tempo do quadro de informação paramétrica adicional dentro da representação de parâmetro (40; 72), caracterizado pelo fato de que a informação paramétrica
10 adicional é usada pelo gerador (32; 62) para gerar os componentes de sinal direto e difuso.

9. Gerador de canais múltiplos (30; 60), de acordo com a reivindicação 8, caracterizado pelo fato de que o gerador (32; 62) é operacional para usar informação paramétrica
15 adicional tendo informação sobre a energia do canal original com relação a outros canais da pluralidade de canais originais.

10. Reconstrutor de canais múltiplos (30; 60), de acordo com qualquer uma das reivindicações anteriores, caracterizado pelo fato de que o modificador de sinal direto (34; 68) é operacional para usar informação sobre uma estrutura
20 temporal do canal original relaciona uma estrutura temporal do canal original com uma estrutura temporal do canal de *downmix* (38; 68).

11. Reconstrutor de canais múltiplos (30; 60), de acordo com qualquer uma das reivindicações anteriores, caracterizado pelo fato de que a informação sobre a estrutura
25 temporal do canal original e a informação sobre a estrutura temporal do canal de *downmix* está tendo uma energia ou uma medição

de amplitude.

12. Reconstructor de canais múltiplos (30; 60), de acordo com qualquer uma das reivindicações anteriores, caracterizado pelo fato de que o modificador de sinal direto (34; 5 68) é ainda operacional para derivar informação temporal de *downmix* na estrutura temporal do canal de *downmix* (38; 68).

13. Reconstructor de canais múltiplos (30; 60), de acordo com a reivindicação 12, caracterizado pelo fato de que o modificador de sinal direto (34; 68) é operacional para derivar 10 informação temporal de *downmix* indicando a energia contida no canal de *downmix* (38; 68) dentro de um intervalo de tempo de comprimento finito ou uma medição de amplitude para o intervalo de tempo de comprimento finito.

14. Reconstructor de canais múltiplos (30; 60), de 15 acordo com as reivindicações 12 ou 13, caracterizado pelo fato de que o modificador de sinal direto (34; 68) é ainda operacional para derivar uma estrutura temporal alvo para o canal de *downmix* reconstruído (38; 68) usando a informação temporal de *downmix* e a informação sobre a estrutura temporal do canal original.

20 15. Reconstructor de canais múltiplos (30; 60), de acordo com qualquer uma das reivindicações de 12 a 14, caracterizado pelo fato de que o modificador de sinal direto (34; 68) é operacional para derivar a informação temporal de *downmix* para uma porção espectral do canal de *downmix* (38; 68) acima do 25 limite inferior espectral.

16. Reconstructor de canais múltiplos (30; 60), de acordo com qualquer uma das reivindicações de 12 a 15, caracterizado pelo fato de que o modificador de sinal direto (34;

68) é ainda operacional para clarear espectralmente o canal de *downmix* (38; 68) e para derivar a informação temporal de *downmix* usando o canal de *downmix* espectralmente clareado (38; 68).

17. Reconstrutor de canais múltiplos (30; 60), de
5 acordo com qualquer uma das reivindicações de 12 a 16, caracterizado pelo fato de que o modificador de sinal direto (34; 68) é ainda operacional para derivar uma representação suavizada do canal de *downmix* (38; 68) e para derivar a informação temporal de *downmix* da representação suavizada do canal de *downmix*.

10 18. Reconstrutor de canais múltiplos (30; 60), de acordo com a reivindicação 17, caracterizado pelo fato de que o modificador de sinal direto (34; 68) é operacional para derivar a representação suavizada pela filtração do canal de *downmix* (38; 68) com um filtro de passagem baixa de primeira ordem.

15 19. Reconstrutor de canais múltiplos (30; 60), de acordo com qualquer uma das reivindicações anteriores, caracterizado pelo fato de que o modificador de sinal direto (34; 68) é ainda operacional para derivar informação sobre uma estrutura temporal de uma combinação do componente de sinal direto
20 e do componente de sinal difuso.

20. Reconstrutor de canais múltiplos (30; 60), de acordo com a reivindicação 19, caracterizado pelo fato de que o modificador de sinal direto (34; 68) é operacional para clarear espectralmente a combinação dos componentes de sinal direto e
25 sinal difuso e para derivar a informação sobre a estrutura temporal da combinação dos componentes do sinal direto e sinal difuso usando os componentes de sinal direto e sinal difuso espectralmente clareados.

21. Recontrutor de canais múltiplos (30; 60), de acordo com as reivindicações 19 ou 20, caracterizado pelo fato de que o modificador de sinal direto (34; 68) é ainda operacional para derivar uma representação suavizada da combinação dos componentes do sinal direto e difuso e para derivar a informação sobre a estrutura temporal da combinação dos componentes do sinal direto e difuso da representação suavizada da combinação dos componentes de sinal direto e de sinal difuso.

22. Recontrutor de canais múltiplos (30; 60), de acordo com a reivindicação 21, caracterizado pelo fato de que o modificador de sinal direto (34; 68) é operacional para derivar a representação suavizada da combinação dos componentes do sinal direto e difuso pela filtração dos componentes do sinal direto e difuso com um filtro de baixa passagem de primeira ordem.

23. Recontrutor de canais múltiplos (30; 60), de acordo com qualquer uma das reivindicações anteriores, caracterizado pelo fato de que o modificador de sinal direto (34; 68) é operacional para usar informação sobre a estrutura temporal do canal original representando uma proporção da energia ou amplitude para um intervalo de tempo de comprimento finito do canal original e a energia ou amplitude para o intervalo de tempo de comprimento finito do canal de *downmix* (38; 68).

24. Recontrutor de canais múltiplos (30; 60), de acordo com qualquer uma das reivindicações anteriores, caracterizado pelo fato de que o modificador de sinal direto (34; 68) é operacional para derivar uma estrutura temporal alvo para o canal de saída reconstruído (50; 76) usando o canal de *downmix* (38; 68) e a informação sobre a estrutura temporal.

25. Recontrutor de canais múltiplos (30; 60), de acordo com a reivindicação 23, caracterizado pelo fato de que o modificador de sinal direto (34; 68) é operacional para modificar o componente de sinal direto, de modo que uma estrutura temporal do canal de saída reconstruído (50; 76) é igual à estrutura temporal alvo dentro de uma faixa de tolerância.

26. Recontrutor de canais múltiplos (30; 60), de acordo com a reivindicação 24, caracterizado pelo fato de que o modificador de sinal direto (34; 68) é operacional para derivar um fator de escalonamento intermediário, o fator de escalonamento intermediário sendo tal que a estrutura temporal do canal de saída reconstruído (50; 76) é igual à estrutura temporal alvo dentro da faixa de tolerância, quando o canal de saída reconstruído (50; 76) é combinado usando os componentes de sinal direto escalonados com o fator de escalonamento intermediário e o componente de sinal difuso escalonado com o fator de escalonamento intermediário.

27. Recontrutor de canais múltiplos (30; 60), de acordo com a reivindicação 25, caracterizado pelo fato de que o modificador de sinal direto (34; 68) é ainda operacional para derivar um fator de escalonamento final usando o fator de escalonamento intermediário e os componentes de sinal direto e difuso de modo que a estrutura temporal do canal de saída reconstruído (50; 76) é igual à estrutura temporal alvo dentro da faixa de tolerância, quando o canal de saída reconstruído (50; 76) é combinado usando o componente de sinal difuso e o componente de sinal direto escalonado usando o fator de escalonamento final.

28. Método para gerar um canal de saída reconstruído (50; 76) usando pelo menos um canal de *downmix* (38;

68) derivado pela execução de *downmix* de uma pluralidade de canais originais e usando uma representação de parâmetro (40; 72), a representação de parâmetro (40; 72) incluindo informação sobre uma estrutura temporal de um canal original, o método caracterizado
5 pelo fato de que compreende: gerar um componente de sinal direto e um componente de sinal difuso para o canal de saída reconstruído (50; 76), baseado no canal de *downmix* (38; 68); modificar o componente de sinal direto usando a representação de parâmetro (40; 72); e combinar o componente de sinal direto modificado (46)
10 e o componente de sinal difuso para obter o canal de saída reconstruído (50; 76).

29. Decodificador de áudio de canais múltiplos para gerar uma reconstrução de um sinal de canais múltiplos caracterizado pelo fato de que usa pelo menos um canal de *downmix*
15 (38; 68) derivado pela execução de *downmix* de uma pluralidade de canais originais e usando uma representação de parâmetro (40; 72), a representação de parâmetro (40; 72) incluindo informação sobre uma estrutura temporal de um canal original, o decodificador de áudio de canais múltiplos compreendendo um reconstrutor de canais
20 múltiplos, de acordo com as reivindicações 1 a 27.

30. Programa de computador com um código de programa para executar o método da reivindicação 28, caracterizado pelo fato de que é executado em um computador.

FIGURA 1

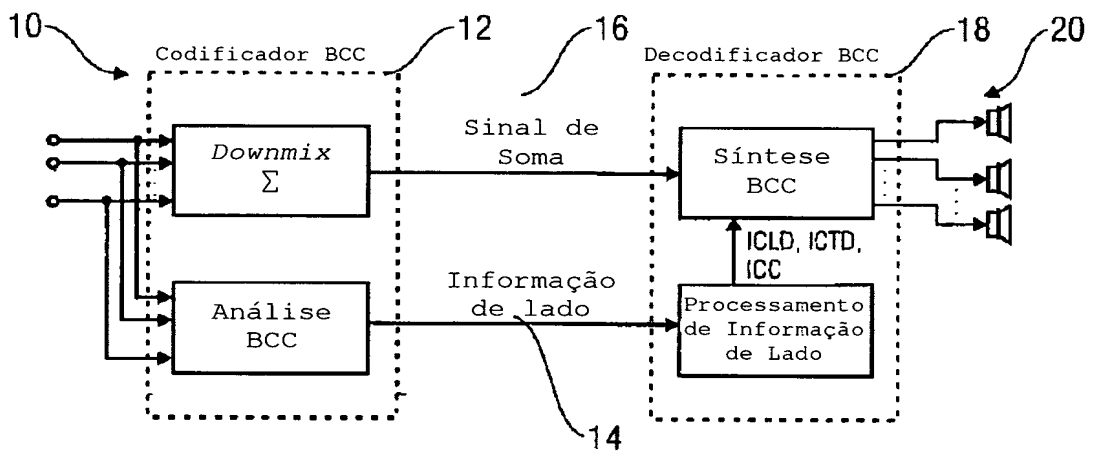


FIGURA 1B

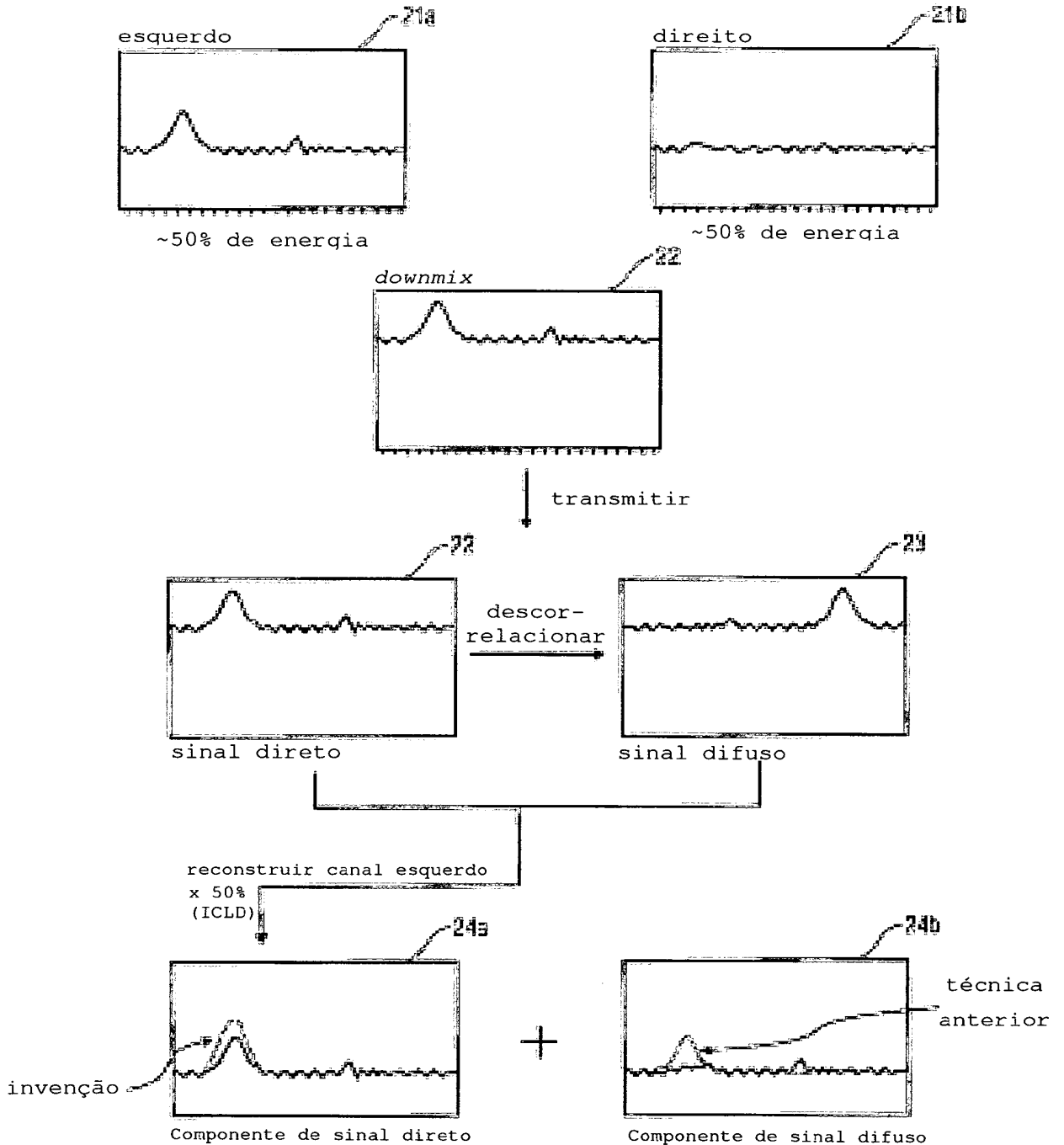


FIGURA 2

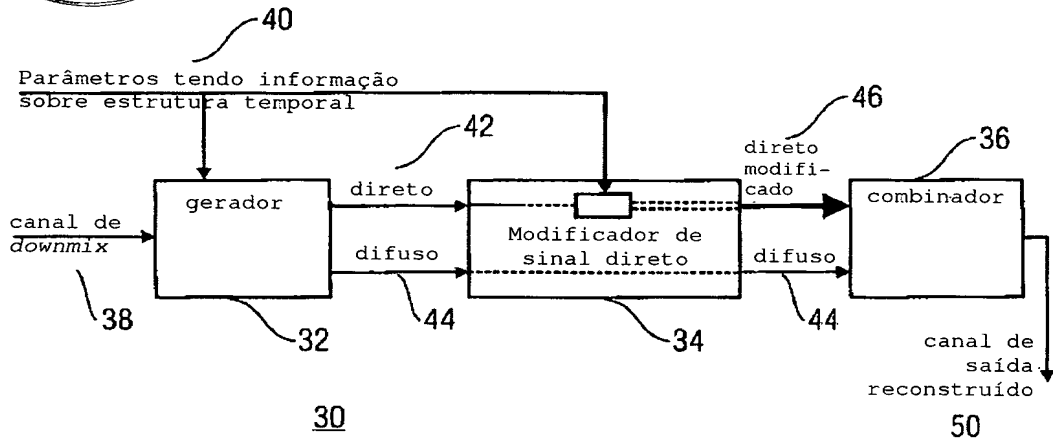


FIGURA 3

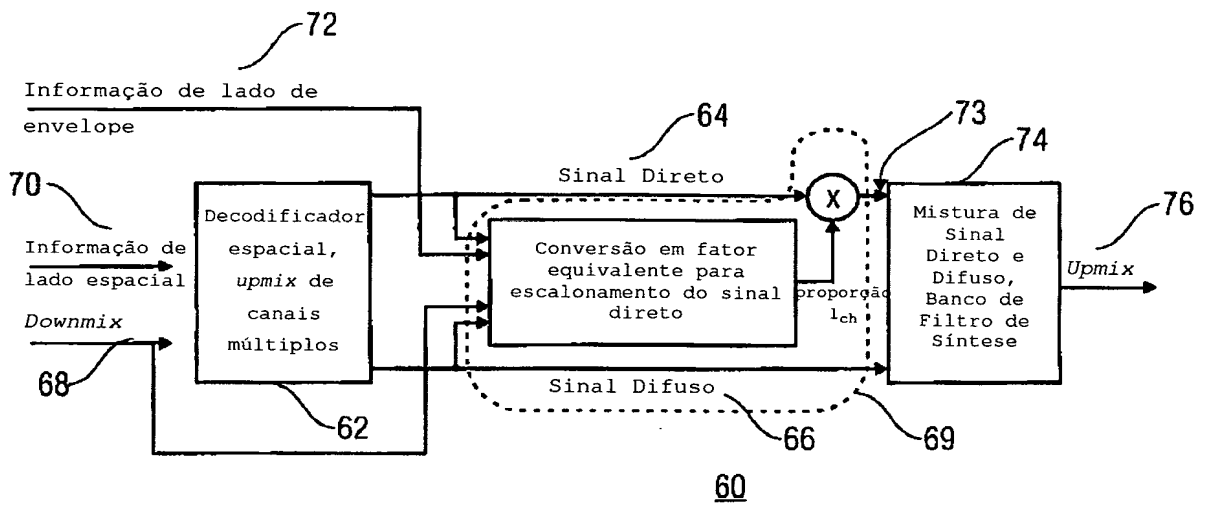


FIGURA 4

k	0	1	2	3	4	5	6	7	8	9	10
$\bar{k}(k)$	1	0	0	1	2	3	4	5	6	7	8
k	11	12	13	14	15	16	17	18	19	20	21
$\bar{k}(k)$	9	10	11	12	13	14	14	15	15	15	16
k	22	23	24	25	26	27	28	29	30	31	32
$\bar{k}(k)$	16	16	16	17	17	17	17	17	18	18	18
k	33	34	35	36	37	38	39	40	41	42	43
$\bar{k}(k)$	18	18	18	18	18	18	18	18	18	19	19
k	44	45	46	47	48	49	50	51	52	53	54
$\bar{k}(k)$	19	19	19	19	19	19	19	19	19	19	19
k	55	56	57	58	59	60	61	62	63	64	65
$\bar{k}(k)$	19	19	19	19	19	19	19	19	19	19	19
k	66	67	68	69	70						
$\bar{k}(k)$	19	19	19	19	19						

FIGURA 5

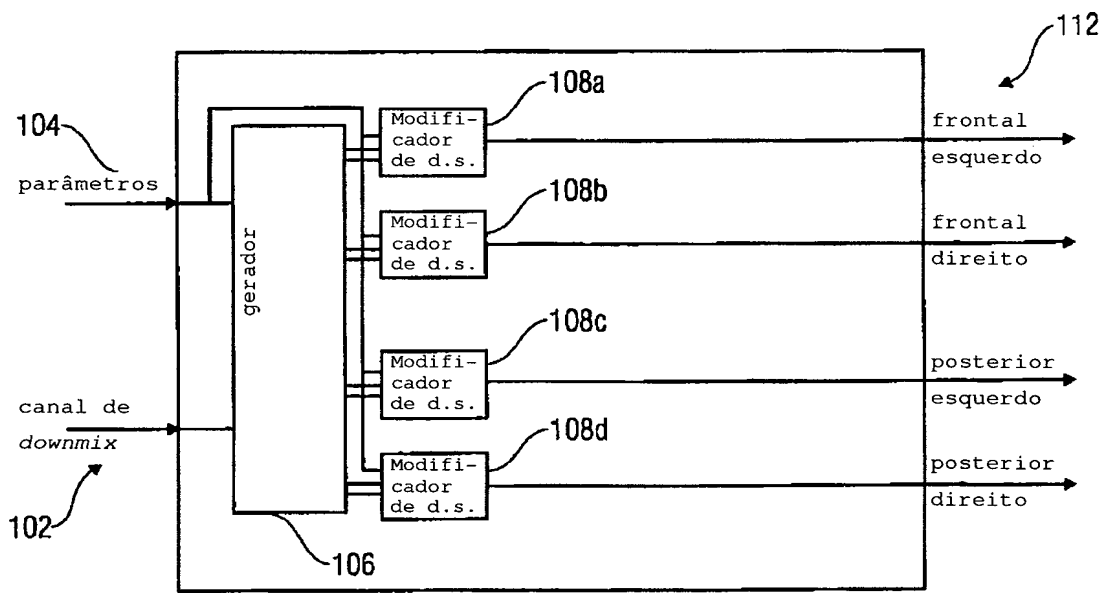
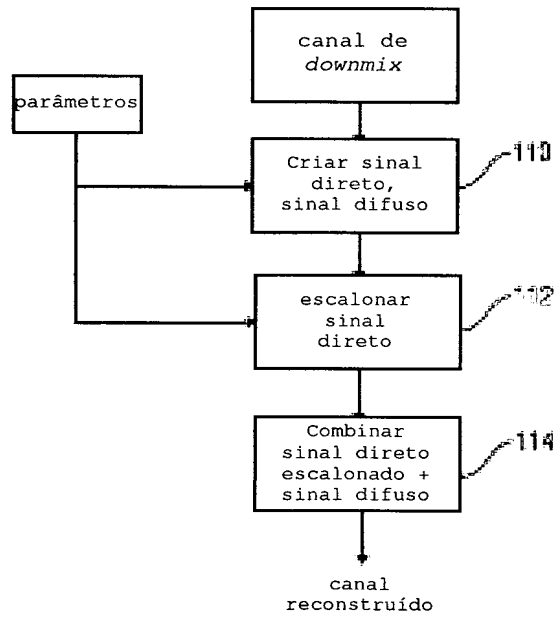


FIGURA 6



RESUMO

"MÉTODO MELHORADO PARA FORMATAÇÃO DE SINAL EM RECONSTRUÇÃO DE ÁUDIO DE CANAIS MÚLTIPLOS"

Canal de saída reconstruído, reconstruído com um
5 reconstrutor de canais múltiplos usando pelo menos um canal de
downmix derivado pela execução de *downmix* de uma pluralidade de
canais originais, e usando uma representação de parâmetro
incluindo informação adicional sobre uma estrutura fina temporal
de um canal original, pode ser gerado usando um gerador (32) para
10 gerar um componente de sinal direto (42), e um componente de sinal
difuso (44) baseado no canal de *downmix* (38) é usado. Apenas o
componente de sinal direto (42) é modificado (34), de modo que a
estrutura fina temporal (40) do canal de saída reconstruído é
adequada a uma estrutura final temporal desejada, indicada pela
15 informação adicional sobre a estrutura fina temporal transmitida.