



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

(11) PI 0806346-0 B1



(22) Data do Depósito: 07/01/2008

(45) Data de Concessão: 29/09/2020

(54) Título: DECODIFICADOR DE ÁUDIO, MÉTODO PARA DECODIFICAR SINAIS DE ÁUDIO, RECEPTOR PARA RECEBER SINAIS DE ÁUDIO, SISTEMA DE COMUNICAÇÃO PARA COMUNICAR SINAIS DE ÁUDIO, MÉTODOS PARA RECEBER SINAIS DE ÁUDIO E PARA TRANSMITIR E RECEBER SINAIS DE ÁUDIO, E DISPOSITIVO DE REPRODUÇÃO DE ÁUDIO

(51) Int.Cl.: G10L 19/26; G10L 19/008; G10L 19/20.

(52) CPC: G10L 19/26; G10L 19/008; G10L 19/20.

(30) Prioridade Unionista: 10/01/2007 EP 07100339.6.

(73) Titular(es): KONINKLIJKE PHILIPS N.V..

(72) Inventor(es): DIRK J. BREEBAART.

(86) Pedido PCT: PCT IB2008050029 de 07/01/2008

(87) Publicação PCT: WO 2008/084427 de 17/07/2008

(85) Data do Início da Fase Nacional: 08/07/2009

(57) Resumo: DECODIFICADOR DE ÁUDIO, MÉTODO PARA DECODIFICAR SINAIS DE ÁUDIO, RECEPTOR PARA RECEBER SINAIS DE ÁUDIO, SISTEMA DE COMUNICAÇÃO PARA COMUNICAR SINAIS DE ÁUDIO, MÉTODOS PARA RECEBER SINAIS DE ÁUDIO E PARA TRANSMITIR E RECEBER SINAIS DE ÁUDIO, E, PRODUTO DE PROGRAMA DE COMPUTADOR. Um decodificador de áudio (100) compreendendo meios de efeito, meios de decodificação e meios de renderização. Os meios de efeito (500) geram sinais de áudio de mistura descendente modificados a partir de sinais de áudio de mistura descendente recebidos; Citados sinais de áudio de mistura descendente recebidos compreendem uma mistura descendente de diversos objetos de áudio. Citados sinais de áudio de mistura descendente modificados são obtidos aplicando efeitos a sinais de áudio estimados correspondentes a objetos de áudio compreendidos nos citados sinais de áudio de mistura descendente recebidos. Citados sinais de áudio estimados são derivados dos sinais de áudio de mistura descendente recebidos com base nos parâmetros recebidos. Citados dados parâmetros recebidos compreendem diversos parâmetros de objeto para cada um dos diversos objetos de áudio. Citados sinais de áudio de mistura descendente modificados com base em um tipo de efeito aplicado, são decodificados pelos meios de decodificação ou renderizados por meios de renderização ou combinados com a saída de (...).

“DECODIFICADOR DE ÁUDIO, MÉTODO PARA DECODIFICAR SINAIS DE ÁUDIO, RECEPTOR PARA RECEBER SINAIS DE ÁUDIO, SISTEMA DE COMUNICAÇÃO PARA COMUNICAR SINAIS DE ÁUDIO, MÉTODOS PARA RECEBER SINAIS DE ÁUDIO E PARA TRANSMITIR
 5 E RECEBER SINAIS DE ÁUDIO, E DISPOSITIVO DE REPRODUÇÃO DE ÁUDIO”

CAMPO TÉCNICO

A invenção relaciona-se a um decodificador de áudio e em particular, porém não exclusivamente, a um decodificador **Surround** MPEG
 10 ou decodificador orientado a objeto.

FUNDAMENTOS TÉCNICOS

Em codificadores áudio espaciais (paramétricos), parâmetros são extraídos dos sinais de áudio, de modo a produzir um número reduzido de sinais de áudio de mistura descendente (por exemplo, apenas um único sinal de
 15 mistura descendente correspondente a uma mistura descendente mono, ou dois sinais de mistura descendente para uma mistura descendente estéreo), e um conjunto correspondente de parâmetros descrevendo as propriedades espaciais do sinal de áudio original. Em decodificadores de áudio espacial (paramétricos), as propriedades espaciais descritas pelos parâmetros espaciais
 20 transmitidos são usadas para recriar um sinal de multicanal espacial, que se parece muito com o sinal de áudio de multicanal original.

Recentemente, técnicas para processamento e manipulação de objetos de áudio individuais no lado da decodificação tem atraído interesse significativo. Por exemplo, dentro da estrutura de trabalho MPEG, um grupo
 25 de trabalho foi iniciado sobre codificação de áudio espacial baseada em objeto. O objetivo deste grupo de trabalho é “explorar nova tecnologia e reutilização de componentes **Surround** MPEG atuais e tecnologias para codificação eficiente de taxa de bit de fontes de som múltiplas ou objetos em um número de canais de mistura descendente e correspondentes parâmetros

espaciais". Em outras palavras, o objetivo é codificar objetos de áudio múltiplos em um conjunto limitado de canais de mistura descendente com parâmetros correspondentes. No lado do decodificador, usuários interagem com o conteúdo, por exemplo, reposicionando os objetos individuais.

5 Tal interação com o conteúdo é facilmente realizada em decodificadores orientados a objeto. Isto é então realizado incluindo uma entrega que segue a decodificação. Citada entrega é combinada com a decodificação, para evitar a necessidade de determinar objetos individuais. A entrega dedicada correntemente disponível compreende posicionamento de
10 objetos, ajuste de volume, ou equalização dos sinais de áudio renderizados.

 Uma desvantagem dos decodificadores orientados a objetos conhecidos com a renderização incorporada é que permitem um conjunto limitado de manipulações de objetos, porque estes não produzem ou operam sobre os objetos individuais. Por outro lado, decodificação explícita dos
15 objetos de áudio individuais é muito dispendiosa e ineficiente.

SUMÁRIO DA INVENÇÃO

 É um objetivo da invenção prover um decodificador melhorado para decodificar objetos de áudio que segue uma faixa mais ampla de manipulações de objetos, sem uma necessidade de decodificar objeto de
20 áudio objetos de áudio individuais para esta finalidade.

 Este objetivo é alcançado por um decodificador de áudio de acordo com a invenção. É suposto que um conjunto de objetos, cada um com sua forma de onda correspondente, tenha sido previamente codificado em um codificador orientado a objeto, o que gera um sinal de áudio de mistura
25 descendente para um unido sinal no caso de um canal único, citado sinal de áudio de mistura descendente sendo uma mistura descendente de diversos objetos de áudio e dados paramétricos correspondentes. Os dados paramétricos compreendem um conjunto de parâmetros de objeto para cada um dos diferentes objetos de áudio. O receptor recebe o citado sinal de áudio

de mistura descendente e os citados dados paramétricos. Este sinal de áudio de mistura descendente é adicionalmente alimentado a meios de efeito que geram sinal de áudio de mistura descendente modificado, aplicando efeitos a estimativas de sinais de áudio correspondendo a objetos de áudio selecionados, compreendidos no sinal de áudio de mistura descendente. Citadas estimativas de sinais de áudio são derivadas com base nos dados paramétricos. O sinal de áudio de mistura descendente modificado é adicionalmente fornecido a meios de decodificação ou meios de renderização, ou combinados com a saída de meios de renderização, dependendo de um tipo de efeito aplicado, por exemplo, um efeito de inserção ou envio. Os meios de decodificação decodificam os objetos de áudio a partir do sinal de áudio de mistura descendente fornecido aos meios de decodificação, citado sinal de áudio de mistura descendente sendo o sinal de áudio de mistura descendente originalmente recebido ou o sinal de áudio de mistura descendente modificado. Citada decodificação é executada com base nos dados paramétricos. Os meios de renderização geram, um sinal de áudio de saída espacial a partir dos objetos de áudio obtidos a partir dos meios de decodificação e opcionalmente a partir dos meios de efeito, dependendo do tipo de efeito aplicado.

A vantagem do decodificador de acordo com a invenção é que, no sentido de aplicar vários tipos de efeito, não é necessário que o objeto, ao qual o efeito deve ser aplicado, esteja disponível. Ao invés disso, a invenção propõe aplicar o efeito aos sinais de áudio estimados correspondendo aos objetos, antes ou em paralelo com a decodificação real. Portanto, decodificação de objeto explícita não é requerida, e a renderização emergente no decodificador é preservada.

Em uma realização, o decodificador compreende adicionalmente meios de modificação para modificar os dados paramétricos quando um envelope espectral ou temporal de um sinal de áudio estimado

correspondente ao objeto ou a diversos objetos, é modificado pelo efeito inserido.

Um exemplo de tal efeito é uma distorção não linear que gera componentes espectrais de alta frequência adicionais, ou um compressor multi-faixa. Se a característica espectral do sinal de áudio modificado, aplicar os parâmetros não modificados compreendidos nos dados paramétricos, conforme recebidos, pode conduzir a artefatos indesejados e possivelmente incômodos. Portanto, adaptar os parâmetros para coincidir com as novas características espectrais ou temporais melhora a qualidade do sinal de áudio resultante renderizado.

Em uma realização, a geração dos sinais de áudio estimados correspondentes a um objeto de áudio ou diversos objetos, compreende escalonamento dependente do tempo/frequência dos sinais de áudio de mistura descendente com base nos parâmetros de potência correspondentes a objetos de áudio, citados parâmetros de potência estando compreendidos nos dados paramétricos recebidos.

A vantagem desta estimativa é que compreende uma multiplicação do sinal de áudio de mistura descendente. Isto torna o processo de estimativa simples e eficiente.

Em uma realização, os meios de decodificação compreendem um decodificador de acordo com o padrão *Surround* MPEG e meios de conversão para converter os dados paramétricos em dados paramétricos de acordo com o padrão *Surround* MPEG.

A vantagem de usar o decodificador *Surround* MPEG é que este tipo de decodificador é usado como uma máquina de renderização para um decodificador orientado a objeto. Neste caso, os parâmetros orientados a objeto são combinados com dados de controle de usuário e convertidos para parâmetros *Surround* MPEG, tais como diferenças de nível e parâmetros de correlação entre canais (pares). Daí, os parâmetros *Surround* MPEG resultam

do efeito combinado de parâmetros orientados a objeto, isto é, informação transmitida, e as propriedades de renderização desejadas, isto é, informação controlável pelo usuário estabelecida no lado do decodificador. Em tal caso, não são requeridos sinais de objeto intermediários.

5 A invenção provê adicionalmente um receptor e um sistema de comunicação, bem como métodos correspondentes.

Em uma realização, os efeitos de inserir e enviar são aplicados simultaneamente. O uso, por exemplo, de efeitos de inserção não exclui o uso de efeitos de envio, e vice-versa.

10 A invenção provê adicionalmente um produto de programa de computador habilitando um dispositivo programável a executar o método de acordo com a invenção.

BREVE DESCRIÇÃO DOS DESENHOS

Estes e outros aspectos da invenção serão aparentes e esclarecidos com referência às realizações mostradas nos desenhos, nos quais:

15 Figura 1A mostra esquematicamente um decodificador orientado a objeto;

Figura 1B mostra esquematicamente um decodificador orientado a objeto de acordo com a invenção;

20 Figura 2 mostra um exemplo de meios de efeito para um efeito de inserção;

Figura 3 mostra meios de modificação para modificar os dados paramétricos quando um envelope espectral de um sinal de áudio estimado correspondente ao objeto ou a diversos objetos é modificado pelo efeito de

25 inserção;

Figura 4 mostra um exemplo de meios de efeito para um efeito de envio;

Figura 5 mostra meios de decodificação, os meios de decodificação compreendendo um decodificador de acordo com padrão

Surround MPEG, e meios de conversão para converter os dados paramétricos em dados paramétricos de acordo com o padrão *Surround* MPEG;

Figura 6 mostra um sistema de transmissão para comunicação de um sinal de áudio de acordo com algumas realizações da invenção.

Através das figuras, os mesmos numerais de referência indicam características similares ou correspondentes. Algumas das características indicadas nos desenhos são tipicamente implementadas em software e como tais representam entidades de software, tais módulos ou objetos de software.

10 DESCRIÇÃO DETALHADA DAS REALIZAÇÕES PREFERIDAS

Figura 1A mostra esquematicamente um decodificador orientado a objeto 100 conforme conhecido, por exemplo, a partir de C.Faller: “Parametric Joint-Coding of Audio Sources”, AES 120th Convention, Paris, France, Preprint 6752, Maio 2006. É suposto que um conjunto de objetos, cada um com sua forma de onda correspondente, foi previamente codificado em um codificador orientado a objeto, o que gera um sinal de áudio de mistura descendente (um único sinal no caso de um canal único, ou dois sinais no caso de dois canais (= estéreo)), citado sinal de áudio de mistura descendente sendo uma mistura descendente de diversos objetos de áudio caracterizados por dados paramétricos correspondentes. Os dados paramétricos compreendem um conjunto de parâmetros de objeto para cada um dos diferentes objetos de áudio. O receptor 200 recebe citado sinal de áudio de mistura descendente e citados dados paramétricos.

O sinal fornecido ao receptor 200 é um único sinal que corresponde ao fluxo de dados de áudio de mistura descendente multiplexados que corresponde ao sinal de áudio de mistura descendente e aos dados paramétricos. A função do receptor é então desmultiplexar os dois fluxos de dados. Se o sinal de áudio de mistura descendente é provido na forma comprimida (tal como a camada 3 MPEG-1) o receptor 200 também executa

descompressão ou decodificação do sinal de áudio comprimido em um sinal de áudio de mistura descendente no domínio do tempo.

Embora a entrada do receptor 200 seja exibida como um caminho único de sinal/dados, esta poderia também compreender caminhos de dados múltiplos para sinais de áudio de mistura descendente separados e/ou dados paramétricos. Conseqüentemente, os sinais de áudio de mistura descendente e os dados paramétricos são fornecidos a meios de decodificação 300 que decodificam os objetos de áudio a partir dos sinais de áudio de mistura descendente, com base nos dados paramétricos. Os objetos de áudio decodificados são adicionalmente fornecidos a meios de renderização 400 para gerar pelo menos um sinal de áudio de saída a partir de objetos de áudio decodificados. Embora os meios de decodificação e meios de renderização sejam conduzidos como unidades separadas, estes freqüentemente são mesclados. Como resultado de tal mesclagem dos meios de decodificação e processamento de renderização, não há necessidade de decodificação explícita de objetos de áudio individuais. Ao invés disso, sinais de áudio renderizados são providos a custo computacional muito mais baixo, e sem perda de qualidade de áudio.

Figura 1B mostra esquematicamente um decodificador orientado a objeto 110 de acordo com a invenção. O receptor 200 recebe o citado sinal de áudio de mistura descendente e os citados dados paramétricos. O sinal de áudio de mistura descendente e os dados paramétricos são adicionalmente fornecidos ao meios de efeito 500 que geram sinal de áudio de mistura descendente modificado, aplicando efeitos a estimativas de sinais de áudio correspondentes a objetos de áudio selecionados compreendidos no sinal de áudio de mistura descendente. Citadas estimativas de sinais de áudio são derivadas com base nos dados paramétricos. O sinal de áudio de mistura descendente modificado é adicionalmente fornecido a meios de decodificação 300, ou meios de renderização 400, ou combinados com a saída de meios de

renderização dependendo de um tipo de efeito aplicado, por exemplo, um efeito de inserção ou envio. Os meios de decodificação 300 decodificam os objetos de áudio a partir do sinal de áudio de mistura descendente fornecido aos meios de decodificação, citando sinal de áudio de mistura descendente sendo o sinal de áudio de mistura descendente recebido originalmente ou o sinal de áudio de mistura descendente modificado. Citada decodificação é efetuada com base nos dados paramétricos. Os meios de renderização 400 geram um sinal de áudio de saída espacial a partir dos objetos de áudio obtidos dos meios de decodificação 300 e opcionalmente a partir dos meios de efeito 400.

Figura 2 mostra um exemplo de meios de efeito 500 para um efeito de inserção. Os sinais de mistura descendente 501 são fornecidos aos meios de efeito 500; estes sinais são fornecidos em paralelo a unidades 511 e 512 que estão compreendidas em meios de estimativas 510. Os meios de estimativas 510 geram os sinais de áudio estimados correspondentes a um objeto ou diversos objetos aos quais o efeito de inserção deve ser aplicado, e o sinal de áudio estimado correspondente aos objetos restantes. A estimativa de sinais de áudio correspondentes a um objeto ou diversos objetos aos quais o efeito de inserção deve ser aplicado, é executada pela unidade 511, enquanto a estimativa do sinal de áudio correspondente aos objetos restantes é executada pela unidade 512. Citada estimativa é baseada nos dados paramétricos 502 que são obtidos do receptor 200. Conseqüentemente, o efeito de inserção são aplicados pelos meios de inserção 530 nos sinais de áudio estimados correspondentes a um objeto ou a diversos objetos aos quais o efeito de inserção deve ser aplicado. Um somador 540 adiciona os sinais de áudio providos a partir dos meios de inserção 530 e o sinal de áudio estimado correspondentes aos objetos restantes, portanto, juntando novamente todos os objetos. O sinal de mistura descendente modificado 503 resultante é adicionalmente fornecido aos meios de decodificação 300 do decodificador

orientado a objeto 110. No restante do texto, todas as vezes que as unidades 200, 300 ou 400 são referidas estão compreendidas em um decodificador orientado a objeto 110.

Os exemplos de efeitos de inserção são entre outros:
5 compressão de faixa dinâmica, geração de distorção (por exemplo, para simular amplificadores de guitarra) ou *vocoder*. Estes tipos de efeitos é aplicado preferivelmente em um conjunto limitado (preferivelmente único) de objetos de áudio.

Figura 3 mostra meios de modificação 600 para modificar os
10 dados paramétricos quando um envelope espectral de um sinal de áudio estimado correspondente ao objeto ou a diversos objetos, é modificado pelo efeito de inserção. As unidades 511 e 512 estão estimando, por exemplo, objetos de áudio individuais, enquanto a unidade 513 estima os objetos de áudio restantes juntos. Os meios de inserção 530 compreendem unidades
15 separadas 531 e 532 que aplicam efeitos de inserção nos sinais estimados obtidos a partir das unidades 511 e 512, respectivamente. Um somador 540 adiciona os sinais de áudio providos a partir dos meios de inserção 530 e o sinal de áudio estimado correspondentes aos objetos restantes, portanto, montando novamente todos os objetos juntos. O sinal de mistura descendente
20 modificado 503 resultante é adicionalmente fornecido aos meios de decodificação 300 do decodificador orientado a objeto 110.

Os efeitos de inserção usados nas unidades 531 e 532 são do mesmo tipo ou diferem. O efeito de inserção usado pela unidade 532 é por exemplo, uma distorção não linear que gera componentes espectrais de alta
25 frequência adicionais, ou um compressor multi-faixa. Se a característica espectral do sinal de áudio modificado tiver mudado, aplicar os parâmetros não modificados compreendidos nos dados paramétricos conforme recebidos nos meios de decodificação 300, pode conduzir a artefatos indesejados e possivelmente incômodos. Portanto, adaptar os dados paramétricos para

coincidir com as novas características espectrais melhora a qualidade do sinal de áudio resultante. Esta adaptação dos dados paramétricos é efetuada na unidade 600. O dados paramétricos adaptados 504 são fornecidos aos meios de decodificação 300 e são usados para decodificar o(s) sinal(is) de mistura descendente modificados 503.

Deveria ser notado que as duas unidades 531 e 532 compreendidas nos meios de inserção 530 são apenas um exemplo. O número de unidades pode variar, dependendo do número de efeitos de inserção a serem aplicados. Adicionalmente, as unidades 531 e 532 podem ser implementadas em hardware ou software.

Figura 4 mostra um exemplo de meios de efeito para um efeito de envio. Os sinais de mistura descendente 501 são fornecidos aos meios de efeito 500, estes sinais são alimentados em paralelo às unidades 511 e 512 que estão compreendidas nos meios de estimativa 510. Os meios de estimativas 510 geram os sinais de áudio estimados correspondentes a um objeto ou diversos objetos aos quais o efeito de envio deve ser aplicado. Citada estimativa é baseada nos dados paramétricos 502 que são obtidos a partir do receptor 200. Conseqüentemente, ganhos são aplicados pelos meios de ganho 560 nos sinais de áudio estimados correspondentes a um objeto ou a diversos objetos obtidos a partir dos meios de estimativa 510. Ganhos, que também poderiam ser referidos como pesos, determinam a quantidade do efeito por objeto ou diversos objetos. Cada uma das unidades 561 e 562 aplica ganho a sinais de áudio individuais obtidos a partir dos meios de estimativa. Cada uma destas unidades pode aplicar vários ganhos. Um somador 540 adiciona os sinais de áudio providos a partir dos meios de ganho 560 e uma unidade 570 aplica o efeito de envio. O sinal resultante 505, também chamado a saída “úmida” é fornecido aos meios de entrega, ou alternativamente é misturado (ou adicionado) à saída dos meios de renderização.

Os exemplos de efeitos de envio são entre outros reverberação,

efeitos de modulação tais como *chorus*, *flanger* ou *phaser*.

Deveria ser notado que as duas unidades 561 e 562 compreendidas nos meios de ganho 560 são apenas um exemplo. O número de unidades pode ser muito dependente do número de sinais correspondendo a objetos de áudio ou diversos objetos de áudio para os quais o nível de efeito de envio deve ser estabelecido.

Os meios de estimativa 510 e os meios de ganho 560 podem ser combinados em uma única etapa de processamento que estima uma combinação ponderada de sinais de objeto múltiplos. Os ganhos 561 e 562 podem ser incorporados nos meios de estimativa 511 e 512, respectivamente. Isto é também descrito nas equações abaixo, onde Q é uma (estimativa de) combinação ponderada de sinais de objeto e é obtida por operação de escalonamento única por *tile* de tempo/frequência.

Os ganhos por objeto ou combinação de objetos podem ser interpretados como “níveis de efeito de envio”. Em várias aplicações, a quantidade de efeito é preferivelmente controlável pelo usuário por objeto. Por exemplo, o usuário pode desejar um dos objetos sem reverberação, outro objeto com uma pequena quantidade de reverberação e ainda outro objeto com a reverberação plena. Em tal exemplo, os ganhos por objeto poderiam ser iguais a 0, 0,5 e 1,0 para cada um dos respectivos objetos.

Em uma realização, a geração de sinais de áudio estimados correspondentes a um objeto de áudio ou a diversos objetos compreende escalonamento dependendo de tempo/frequência dos sinais de áudio de mistura descendente com base nos parâmetros de potência correspondendo a objetos de áudio, citados parâmetros de potência estando compreendidos nos dados paramétricos.

Esta realização é explicada para o exemplo a seguir. No codificador, I sinais de objeto $s_i[n]$, $i = 0, \dots, I-1$, com n sendo o índice de amostra, são misturados descendentemente para criar um sinal de mistura

descendente $x[n]$, pelo somatório dos sinais de mistura descendente:

$$x[n] = \sum_i s_i[n]$$

O sinal de mistura descendente é acompanhado por parâmetros orientados a objeto que descrevem a potência de sinal (relativa) de cada objeto dentro de *tiles* de tempo/frequência do sinal de mistura descendente $x[n]$. Os sinais de objeto $s_i[n]$ são por exemplo, primeiramente colocados em janelas usando janelas de análise superpostas $w[n]$:

$$s_i[n, m] = s_i[n + mL/2]w[n],$$

Com L sendo a extensão da janela e, por exemplo, $L/2$ o tamanho de transposição correspondente (supondo 50% de superposição) e m o índice de janela. Uma forma típica da janela de análise é uma janela de Hanning:

$$w[n] = \text{sen}\left(\frac{\pi n}{L}\right).$$

Os sinais segmentados resultantes $s_i[n, m]$ são subseqüentemente transformados para o domínio da frequência, usando uma FFT.

$$S_i[k, m] = \sum_n s_i[n, m] e^{-2\pi j k n / L}$$

Com k sendo um índice bin FFT. Os índices bin FFT k são subseqüentemente agrupados em bandas de parâmetro b . Em outras palavras, cada banda de parâmetro b correspondente a um conjunto de índices bin de frequência adjacente k . Para cada banda de parâmetro b e cada segmento m de cada sinal de objeto $S_i[k, m]$, um valor de potência $\sigma_i^2[b, m]$ é computado:

$$\sigma_i^2[b, m] = \frac{\sum_{k=k(b)}^{k=(b+1)-1} S_i[k, m] S_i^*[k, m]}{k(b+1) - k(b)},$$

com $(*)$ sendo o operador de conjugação complexo. Estes

parâmetros $\sigma_i^2[b, m]$ estão compreendidos nos dados paramétricos (preferivelmente quantizados no domínio logarítmico).

O processo de estimativa de um objeto ou diversos objetos no decodificador de áudio orientado a objeto compreende escalonamento dependente do tempo/frequência do sinal de áudio de mistura descendente. Um sinal de mistura descendente discreto no tempo $x[n]$ com n sendo o mesmo índice, é dividido em *tiles* de tempo/frequência $X[k, m]$ com k sendo o índice de frequência e m um índice de quadro (segmento temporal). Isto é obtido, por exemplo, colocando em janelas o sinal $x[n]$ com uma janela de análise $w[n]$:

$$x[n, m] = x[n + mL/2]w[n],$$

Sendo L a extensão da janela e $L/2$ o tamanho de transposição correspondente. Neste caso, uma janela de análise preferida é dada pela raiz quadrada da janela de Hanning:

$$w[n] = \sqrt{\sin\left(\frac{\pi n}{L}\right)}$$

Subseqüentemente, o sinal em janelas $x[n, m]$ é transformado para o domínio da frequência usando uma FFT:

$$X[k, m] = \sum_n x[n, m] e^{-2\pi jkn/L}$$

As componentes no domínio da frequência de $X[k, m]$ são subseqüentemente agrupadas nas assim chamadas bandas de parâmetro b ($b = 0, \dots, B-1$). Estas bandas de parâmetro coincidem com as bandas de parâmetro no codificador. A estimativa do lado do decodificador $\hat{S}_i[k, m]$ do segmento m do objeto i é dada por:

$$\hat{S}_i[k, m] = X[k, m] \sqrt{\frac{\sigma_i^2[b(k), m]}{\sum_i \sigma_i^2[b(k), m]}}$$

Com $b(k)$ sendo a banda de parâmetro que foi associada ao índice de frequência k .

Uma combinação ponderada Q dos sinais de objeto S_i com pesos g_i é dada por:

$$Q[k, m] = \sum_i g_i S_i[k, m]$$

5 No decodificador orientado a objeto, Q pode ser estimado de acordo com:

$$\hat{Q}[k, m] = \sum_i g_i \hat{S}_i[k, m] = X[k, m] \sqrt{\frac{g_i^2 \sigma_i^2[b(k), m]}{\sum_i \sigma_i^2[b(k), m]}}$$

Em outras palavras, um sinal de objeto ou qualquer combinação linear de diversos sinais de objeto de áudio podem ser estimados no decodificador de áudio orientado a objeto proposto, por um escalonamento
10 dependente de tempo/frequência do sinal de mistura descendente $X[k, m]$.

No sentido de resultar nos sinais de saída no domínio do tempo, cada sinal de objeto estimado é transformado para o domínio do tempo (usando uma FFT inversa) multiplicado por uma janela de síntese (idêntica à janela de análise) e combinado com quadros prévios usando superposição-
15 som.

Em uma realização, a geração dos sinais de áudio estimados compreende ponderar um objeto ou uma combinação de diversos objetos por meio de escalonamento dependente de tempo/frequência dos sinais de áudio de mistura descendente, com base nos parâmetros de potência
20 correspondentes a objetos de áudio, citados parâmetros de potência sendo compreendidos nos dados paramétricos recebidos.

Deveria ser notado que uma unidade de efeito de envio poderia ter mais sinais de saída que sinais de entrada. Por exemplo, no caso, uma unidade de reverberação estéreo ou multicanal possui um sinal de entrada

mono.

Em uma realização, o sinal misturado descendentemente e os dados paramétricos estão de acordo com o padrão *Surround* MPEG. O decodificador *Surround* MPEG existente seguinte à funcionalidade de
5 decodificação também funciona como um dispositivo de renderização. Em tal caso, não são requeridos sinais de áudio intermediários correspondentes aos objetos de decodificação. A decodificação de objeto e renderização são combinadas em um único dispositivo.

Figura 5 mostra meios de decodificação, os meios de
10 decodificação 300 compreendem um decodificador 320 de acordo com o padrão *Surround* MPEG e meios de conversão 310 para converter dados paramétricos em dados paramétricos de acordo com o padrão *Surround* MPEG. O(s) sinal(is) 508 correspondente(s) ao(s) sinal(is) de mistura descendente 501 ou sinal(is) de mistura descendente modificado(s)
15 503, quando os efeitos de inserção são aplicados, são alimentados ao decodificador *Surround* MPEG 320. Os meios de conversão baseados nos dados paramétricos 506 e dados de controle de usuário 507 convertem os dados paramétricos em dados paramétricos de acordo com o padrão *Surround* MPEG. Os dados paramétricos 506 são os dados paramétricos 502
20 ou os dados paramétricos modificados 504, quando o envelope espectral de um sinal de áudio estimado correspondente ao objeto ou a diversos objetos, é modificado pelo efeito de inserção. Os dados de controle de usuário 507 para o objeto ou diversos objetos, é modificado pelo efeito de inserção. Os dados de controle de usuário 507 podem, por exemplo, indicar a posição espacial
25 desejada de um ou diversos objetos de áudio.

De acordo com uma das realizações, o método compreende as etapas de receber pelo menos um sinal de áudio de mistura descendente e dados paramétricos, gerando sinais de áudio de mistura descendente modificados, decodificando os objetos de áudio a partir dos sinais de áudio de

mistura descendente, e gerando pelo menos um sinal de áudio de saída a partir dos objetos de áudio decodificados. No método, cada sinal de áudio de mistura descendente compreende uma mistura descendente de diversos objetos de áudio. Os dados paramétricos compreendem diversos parâmetros de objeto para cada um dos diversos objetos de áudio. Os sinais de áudio de mistura descendente modificados são obtidos aplicando efeitos a sinais de áudio estimados correspondendo a objetos de áudio compreendidos nos citados sinais de áudio de mistura descendente. Os sinais de áudio estimados são derivados dos sinais de áudio de mistura descendente, com base nos dados paramétricos. Os sinais de áudio de mistura descendente modificados baseados em um tipo do efeito aplicado são decodificados pelos meios de decodificação 300 ou renderizados pelos meios de renderização 400. A etapa de decodificação é executada pelos meios de decodificação 300 para os sinais de áudio de mistura descendente ou sinais de áudio de mistura descendente modificados, com base nos dados paramétricos.

A última etapa de gerar pelo menos um sinal de áudio de saída a partir dos objetos de áudio decodificados, que pode ser chamada uma etapa de renderização, pode ser combinada com a etapa de decodificação em uma etapa de processamento.

Em uma realização, um receptor para receber sinais de áudio compreende: um elemento de recepção, meios de efeito, meios de decodificação e meios de renderização. O elemento receptor recebe, de um transmissor, pelo menos um sinal de áudio de mistura descendente e dados paramétricos. Cada sinal de áudio de mistura descendente compreende uma mistura descendente de diversos objetos de áudio. Os dados paramétricos compreendem diversos parâmetros de objeto para cada um dos diversos objetos de áudio.

Os meios de efeito geram sinais de áudio de mistura descendente modificados. Estes sinais de áudio de mistura descendente

modificados são obtidos aplicando efeitos a sinais de áudio estimados correspondentes a objetos de áudio compreendidos nos citados sinais de áudio de mistura descendente. Os sinais de áudio estimados são derivados dos sinais de áudio de mistura descendente com base nos dados paramétricos. Os sinais de áudio de mistura descendente modificados, baseados em um tipo do efeito aplicado, são decodificados pelos meios de decodificação ou renderizados pelos meios de renderização.

Os meios de decodificação decodificam os objetos de áudio a partir dos sinais de áudio de mistura descendente ou dos sinais de áudio de mistura descendente modificados, com base nos dados paramétricos. Os meios de renderização geram pelo menos um sinal de áudio de saída a partir dos objetos de áudio decodificados.

Figura 6 mostra um sistema de transmissão para um sinal de áudio de acordo com algumas realizações da invenção. O sistema de transmissão compreende um transmissor 700, que é acoplado a um receptor 900 através de uma rede 800. A rede 800 poderia ser por exemplo, a Internet.

O transmissor 700 é por exemplo, um dispositivo de gravação de sinal e o receptor 900 é por exemplo, um dispositivo reprodutor de sinal. No exemplo específico, quando uma função de gravação de sinal é suportada, o transmissor 700 compreende meios 710 para receber diversos objetos de áudio. Conseqüentemente, estes objetos são codificados pelos meios de codificação 720 para codificar os diversos objetos de áudio pelo menos em um sinal de áudio de mistura descendente e dados paramétricos. Uma realização de tais meios de codificação é dada em Faller, C., "Parametric Joint-Coding of Audio Sources", Proc 120th AES Convention, Paris, France, Maio 2006. Cada sinal de áudio de mistura descendente compreende uma mistura descendente de diversos objetos de áudio. Citados dados paramétricos compreendem diversos parâmetros de objeto para cada um dos diversos objetos de áudio. Os objetos de áudio codificados são transmitidos a um

receptor 900 por meios 730 para transmitir sinais de áudio de mistura descendente e os dados paramétricos. Citados meios 730 possuem uma interface com a rede 800, e podem transmitir os sinais de áudio de mistura descendente através da rede 800.

5 O receptor 900 compreende um elemento receptor 910 para receber do transmissor 700 pelo menos um sinal de áudio de mistura descendente e dados paramétricos. Cada sinal de áudio de mistura descendente compreende uma mistura descendente de diversos objetos de áudio. Citados dados paramétricos compreendem diversos parâmetros de
10 objeto para cada um dos diversos objetos de áudio. Os meios de efeito 920 gerar sinais de áudio de mistura descendente modificados. Citados sinais de áudio de mistura descendente modificados são obtidos, aplicando efeitos a sinais de áudio estimados correspondentes a objetos de áudio compreendidos nos citados sinais de áudio de mistura descendente. Citados sinais de áudio
15 estimados são derivados dos sinais de áudio de mistura descendente, com base nos dados paramétricos. Citados sinais de áudio de mistura descendente modificados baseados em um tipo de efeito aplicado são decodificados pelos meios de decodificação, ou renderizados por meios de renderização, ou combinados com a saída de meios de renderização. Os meios de
20 decodificação decodificam os objetos de áudio a partir dos sinais de áudio de mistura descendente ou dos sinais de áudio de mistura descendente modificados, com base nos dados paramétricos. Os meios de renderização geram pelo menos um sinal de áudio de saída a partir dos objetos de áudio decodificados.

25 Em uma realização, os efeitos de inserção e envio são aplicados simultaneamente.

Em uma realização, os efeitos são aplicados em resposta a entrada de usuário. O usuário pode, por meio, por exemplo, de um botão controle deslizante, botão ou interface de uso gráfico, ajustar os efeitos de

acordo com as próprias preferências.

Deveria ser notado que as realizações acima mencionadas ilustram ao invés de limitar a invenção, e aqueles especialistas na técnica serão capazes de projetar muitas realizações alternativas, sem se afastar do escopo das reivindicações anexas.

Nos desenhos que a acompanham, quaisquer sinais de referência colocados entre parênteses não serão considerados como limitando a reivindicação. A palavra “compreendendo” não exclui a presença de elementos ou etapas diferentes daqueles listados em uma reivindicação. A palavra “um” ou “uma” precedendo um elemento não exclui a presença de diversos de tais elementos. A invenção pode ser implementada por meio de hardware compreendendo vários elementos distintos, e por meio de um computador adequadamente programado.

REIVINDICAÇÕES

1. Decodificador de áudio (100), caracterizado por compreender:

meios de efeito (500) para gerar sinais de áudio de mistura descendente modificados a partir de sinais de áudio de mistura descendente recebidos, citados sinais de áudio de mistura descendente compreendendo uma mistura descendente de diversos objetos de áudio, citados sinais de áudio de mistura descendente modificados obtidos aplicando efeitos a sinais de áudio estimados correspondentes a objetos de áudio compreendidos nos citados sinais de áudio de mistura descendente recebidos, citados sinais de áudio estimados sendo derivados dos sinais de áudio de mistura descendente recebidos com base em dados paramétricos recebidos, citados dados paramétricos recebidos compreendendo diversos parâmetros de objeto para cada um dos diversos objetos de áudio, citados sinais de áudio de mistura descendente modificados com base em um tipo de efeito aplicado, sendo decodificados pelos meios de decodificação ou renderizados por meios de renderização ou combinados com a saída de meios de renderização;

os meios de decodificação (300) sendo arranjos para decodificar os objetos de áudio a partir de sinais de áudio de mistura descendente ou sinais de áudio de mistura descendente modificados com base nos dados paramétricos;

os meios de renderização (400) sendo arranjos para gerar pelo menos um sinal de áudio de saída a partir dos objetos de áudio decodificados.

2. Decodificador de acordo com a reivindicação 1, caracterizado pelos meios de efeito (500) serem arranjos para prover um efeito de inserção e compreenderem:

meios de estimativa (510) para gerar os sinais de áudio estimados correspondentes a um objeto ou diversos objetos aos quais o efeito de

inserção deve ser aplicado, e gerar o sinal de áudio estimado correspondente aos objetos restantes;

meios de inserção (530) para aplicar o efeito de inserção aos sinais de áudio estimados correspondentes a um objeto ou diversos objetos aos
5 quais o efeito de inserção deve ser aplicado;

somador (540) para adicionar os sinais de áudio providos a partir dos meios de inserção e o sinal de áudio estimado correspondente aos objetos restantes.

3. Decodificador de acordo com a reivindicação 2, caracterizado
10 pelo decodificador compreender adicionalmente meios de modificação (600) para modificar os dados paramétricos quando um envelope espectral ou temporal de um sinal de áudio estimado correspondente ao objeto ou diversos objetos, é modificado pelo efeito de inserção.

4. Decodificador de acordo com a reivindicação 1, caracterizado
15 pelos meios de efeito serem arrançados para prover um efeito de envio e compreenderem:

meios de estimativas (510) para gerar os sinais de áudio estimados correspondentes a um objeto ou diversos objetos aos quais o efeito de envio deve ser aplicado;

20 meios de ganho (560) para determinar uma quantidade do efeito de envio para os sinais de áudio estimados correspondentes ao objeto ou os diversos objetos aos quais o efeito de envio deve ser aplicado;

somador (540) para adicionar os sinais de áudio obtidos a partir dos meios de ganho;

25 meios de envio (570) para aplicar o efeito de envio aos sinais de áudio obtidos a partir do somador.

5. Decodificador de acordo com a reivindicação 1, caracterizado pela geração dos sinais de áudio estimados correspondentes a um objeto de áudio

ou diversos objetos, compreenderem escalonamento dependente de tempo/frequência dos sinais de áudio de mistura descendente com base nos parâmetros de potência correspondentes aos objetos de áudio, citados parâmetros de potência estando compreendidos nos dados paramétricos.

5 6. Decodificador de acordo com a reivindicação 5, caracterizado pela geração dos sinais de áudio estimados compreender ponderação de um objeto ou uma combinação de diversos objetos, por meio de escalonamento dependente de tempo/frequência dos sinais de áudio de mistura descendente, com base nos parâmetros de potência correspondentes a objetos de áudio, citados
10 parâmetros de potência estando compreendidos nos dados paramétricos recebidos.

7. Decodificador de acordo com a reivindicação 1, caracterizado pelo sinal misturado descendendemente e os dados paramétricos estarem de acordo com o padrão **Surround** MPEG.

15 8. Decodificador de acordo com a reivindicação 7, caracterizado pelos meios de decodificação (300) compreenderem um decodificador (320) de acordo com o padrão **Surround** MPEG e meios de conversão (310) para converter os dados paramétricos em dados paramétricos de acordo com o padrão **Surround** MPEG.

20 9. Método para decodificar sinais de áudio, caracterizado por compreender:

receber pelo menos um sinal de áudio de mistura descendente e dados paramétricos, cada sinal de áudio de mistura descendente compreendendo uma mistura descendente de diversos objetos de áudio, citados dados
25 paramétricos compreendendo diversos parâmetros de objeto para cada um dos diversos objetos de áudio;

gerar sinais de áudio de mistura descendente modificados;
citados sinais de áudio de mistura descendente obtidos aplicando efeitos a sinais

de áudio estimados correspondentes a objetos de áudio compreendidos nos citados sinais de áudio de mistura descendente, citados sinais de áudio estimados sendo derivados dos sinais de áudio de mistura descendente com base nos dados paramétricos, citados sinais de áudio de mistura descendente modificados
 5 baseados em um tipo de efeito aplicado, sendo decodificados por meios de decodificação ou renderizados por meios de renderização ou combinados com a saída dos meios de renderização;

decodificar os objetos de áudio a partir dos sinais de áudio de mistura descendente ou dos sinais de áudio de mistura descendente modificados,
 10 com base nos dados paramétricos;

gerar pelo menos um sinal de áudio de saída a partir dos objetos de áudio decodificados.

10. Método, de acordo com a reivindicação 9, caracterizado pelos efeitos de inserção e envio serem aplicados simultaneamente.

11. Método, de acordo com a reivindicação 9, caracterizado pelos efeitos serem aplicados em resposta a entrada de usuário.

12. Receptor para receber sinais de áudio, caracterizado por compreender:

decodificador de áudio compreendendo:

20 meios de efeito (500) para gerar sinais de áudio de mistura descendente modificados a partir de sinais de áudio de mistura descendente recebidos, citados sinais de áudio de mistura descendente compreendendo uma mistura descendente de diversos objetos de áudio, citados sinais de áudio de mistura descendente modificados obtidos aplicando efeitos a sinais de áudio
 25 estimados correspondentes a objetos de áudio compreendidos nos citados sinais de áudio de mistura descendente recebidos, citados sinais de áudio estimados sendo derivados dos sinais de áudio de mistura descendente recebidos com base em dados paramétricos recebidos, citados dados paramétricos recebidos

compreendendo diversos parâmetros de objeto para cada um dos diversos objetos de áudio, citados sinais de áudio de mistura descendente modificados com base em um tipo de efeito aplicado, sendo decodificados pelos meios de decodificação ou renderizados por meios de renderização ou combinados com a saída de meios de renderização;

os meios de decodificação (300) sendo arranjos para decodificar os objetos de áudio a partir de sinais de áudio de mistura descendente ou sinais de áudio de mistura descendente modificados com base nos dados paramétricos;

os meios de renderização (400) sendo arranjos para gerar pelo menos um sinal de áudio de saída a partir dos objetos de áudio decodificados; e um elemento receptor (200) para receber, de um transmissor, pelo menos um sinal de áudio de mistura descendente e dados paramétricos, cada sinal de áudio de mistura descendente compreendendo uma mistura descendente de diversos objetos de áudio, citados dados paramétricos compreendendo diversos parâmetros de objeto para cada um dos diversos objetos de áudio, o elemento receptor sendo acoplado aos meios de efeito (500) e meios de decodificação (300).

13. Sistema de comunicação para comunicar sinais de áudio, caracterizado por compreender:

transmissor (700) compreendendo: meios (710) para receber diversos objetos de áudio, meios de codificação (720) para codificar os diversos objetos de áudio pelo menos em um sinal de áudio de mistura descendente e dados paramétricos, cada sinal de áudio de mistura descendente compreendendo uma mistura descendente de diversos objetos de áudio, citados dados paramétricos compreendendo diversos parâmetros de objeto para cada um dos diversos objetos de áudio, e

meios (730) para transmitir sinais de áudio de mistura

descendente e os dados paramétricos a um receptor; e

o receptor (900) compreendendo:

um elemento receptor (200) para receber, do citado transmissor, pelo menos um sinal de áudio de mistura descendente e dados paramétricos, cada

5 sinal de áudio de mistura descendente compreendendo uma mistura descendente de diversos objetos de áudio, citados dados paramétricos compreendendo diversos parâmetros de objeto para cada um dos diversos objetos de áudio, e

um elemento de decodificação compreendendo meios de decodificação e meios de efeito, o elemento receptor sendo acoplado a meios de

10 efeito e a meios de decodificação.

14. Método para receber sinais de áudio, caracterizado por compreender:

receber de um transmissor pelo menos um sinal de áudio de mistura descendente e dados paramétricos, cada sinal de áudio de mistura
15 descendente compreendendo uma mistura descendente de diversos objetos de áudio, citados dados paramétricos compreendendo diversos parâmetros de objeto para cada um dos diversos objetos de áudio;

gerar sinais de áudio de mistura descendente modificados; citados sinais de áudio de mistura descendente modificados obtidos aplicando
20 efeitos a sinais de áudio estimados correspondentes a objetos de áudio compreendidos nos citados sinais de áudio de mistura descendente, citados sinais de áudio estimados sendo derivados dos sinais de áudio de mistura descendente com base nos dados paramétricos, citados sinais de áudio de mistura descendente modificados baseados em um tipo de efeito aplicado, sendo decodificados por
25 meios de decodificação ou renderizados por meios de renderização ou combinados com a saída dos meios de renderização;

decodificar os objetos de áudio a partir dos sinais de áudio de mistura descendente ou dos sinais de áudio de mistura descendente modificados,

com base nos dados paramétricos,

meios de renderização para gerar pelo menos um sinal de áudio de saída a partir dos objetos de áudio decodificados.

15 15. Método para transmitir e receber sinais de áudio, caracterizado por compreender:

em um transmissor, efetuar as etapas de: receber diversos objetos de áudio,

codificar os diversos objetos de áudio pelo menos em um sinal de áudio de mistura descendente e dados paramétricos, cada sinal de áudio de
10 mistura descendente compreendendo uma mistura descendente de diversos objetos de áudio, citados dados paramétricos compreendendo diversos parâmetros de objeto para cada um dos diversos objetos de áudio, e

transmitir sinais de áudio de mistura descendente e os dados paramétricos a um receptor; e

15 no receptor efetuar as etapas de:

receber do transmissor pelo menos um sinal de áudio de mistura descendente e dados paramétricos, cada sinal de áudio de mistura descendente compreendendo uma mistura descendente de diversos objetos de áudio, citados dados paramétricos compreendendo diversos parâmetros de objeto para cada um
20 dos diversos objetos de áudio,

gerar sinais de áudio de mistura descendente modificados; citados sinais de áudio de mistura descendente modificados obtidos, aplicando efeitos a sinais de áudio estimados correspondentes a objetos de áudio compreendidos nos citados sinais de áudio de mistura descendente, citados sinais
25 de áudio estimados sendo derivados dos sinais de áudio de mistura descendente, com base nos dados paramétricos, citados sinais de áudio de mistura descendente modificados baseados em um tipo de efeito aplicado sendo decodificados pelos meios de decodificação, ou renderizados por meios de renderização, ou

combinados com a saída de meios de renderização;

decodificar os objetos de áudio a partir dos sinais de áudio de mistura descendente ou dos sinais de áudio de mistura descendente modificados, com base nos dados paramétricos,

5 gerar pelo menos um sinal de áudio de saída a partir dos objetos de áudio decodificados.

16. Dispositivo de reprodução de áudio, caracterizado por compreender um decodificador de áudio que compreende:

 meios de efeito (500) para gerar sinais de áudio de mistura
10 descendente modificados a partir de sinais de áudio de mistura descendente recebidos, citados sinais de áudio de mistura descendente compreendendo uma mistura descendente de diversos objetos de áudio, citados sinais de áudio de mistura descendente modificados obtidos aplicando efeitos a sinais de áudio estimados correspondentes a objetos de áudio compreendidos nos citados sinais
15 de áudio de mistura descendente recebidos, citados sinais de áudio estimados sendo derivados dos sinais de áudio de mistura descendente recebidos com base em dados paramétricos recebidos, citados dados paramétricos recebidos compreendendo diversos parâmetros de objeto para cada um dos diversos objetos de áudio, citados sinais de áudio de mistura descendente modificados
20 com base em um tipo de efeito aplicado, sendo decodificados pelos meios de decodificação ou renderizados por meios de renderização ou combinados com a saída de meios de renderização;

 meios de decodificação (300) sendo arrançados para decodificar os objetos de áudio a partir de sinais de áudio de mistura descendente ou sinais
25 de áudio de mistura descendente modificados com base nos dados paramétricos;

 meios de renderização (400) sendo arrançados para gerar pelo menos um sinal de áudio de saída a partir dos objetos de áudio decodificados.

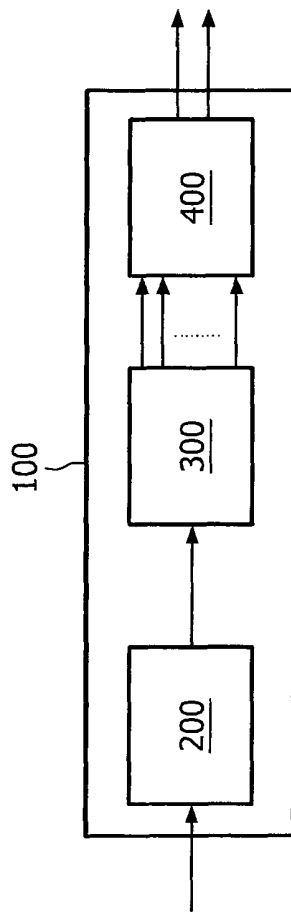


FIG. 1A

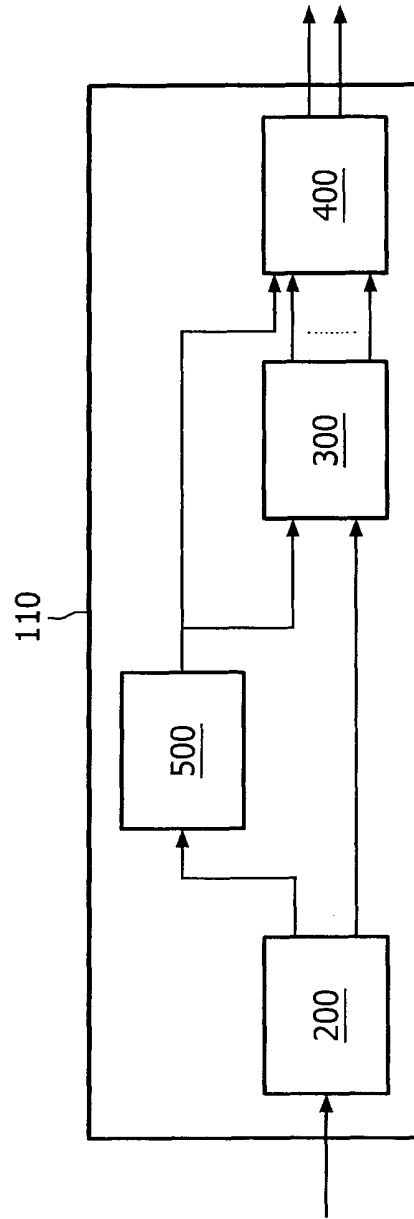


FIG. 1B

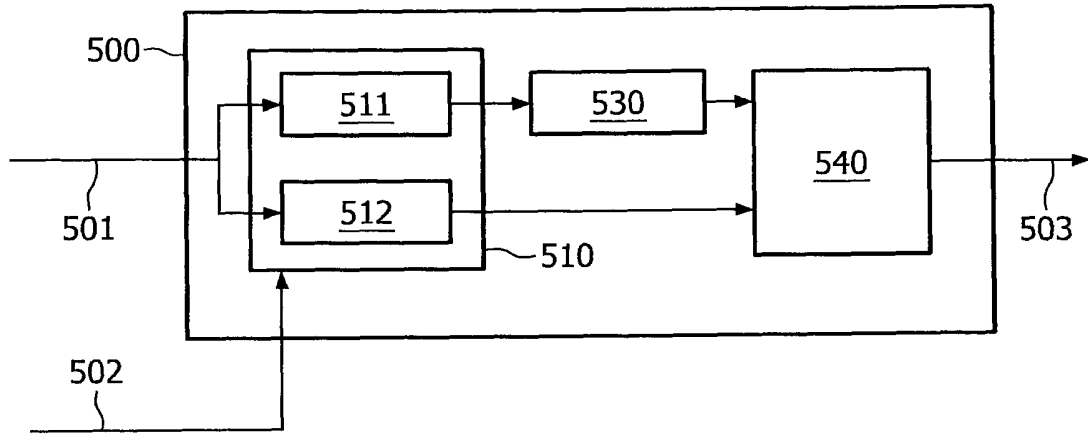


FIG. 2

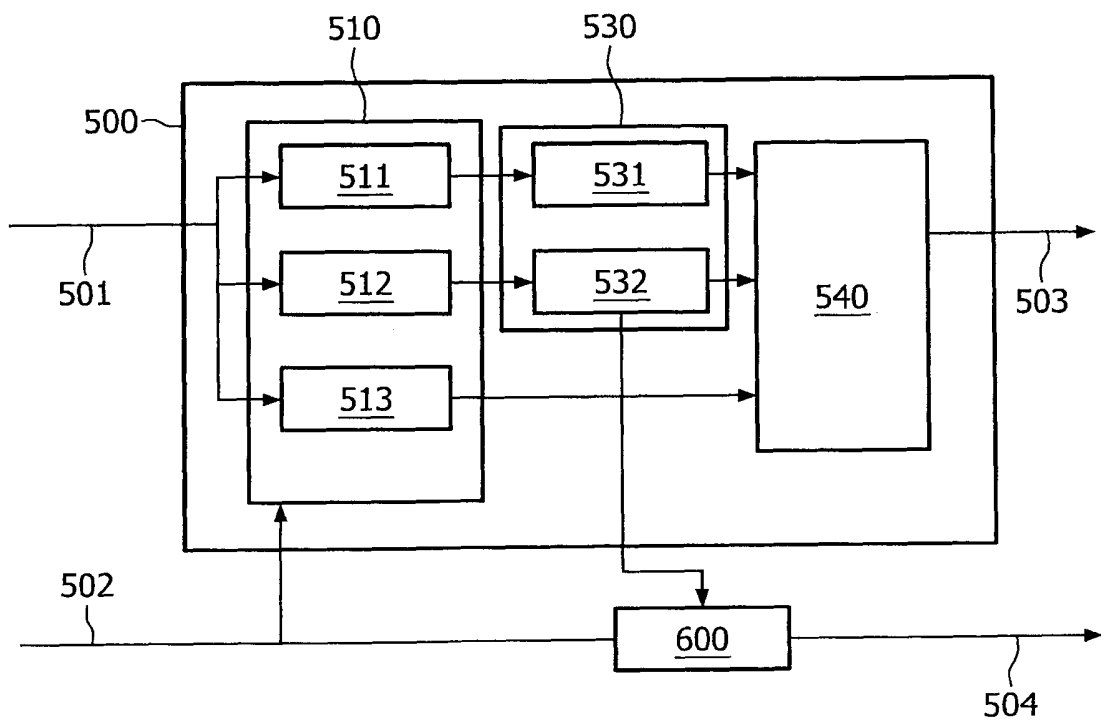


FIG. 3

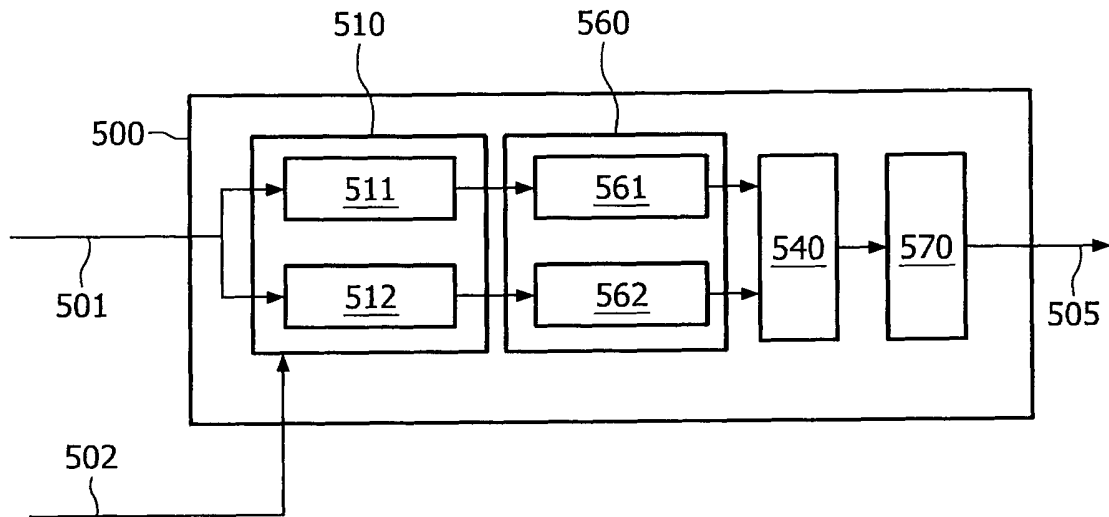


FIG. 4

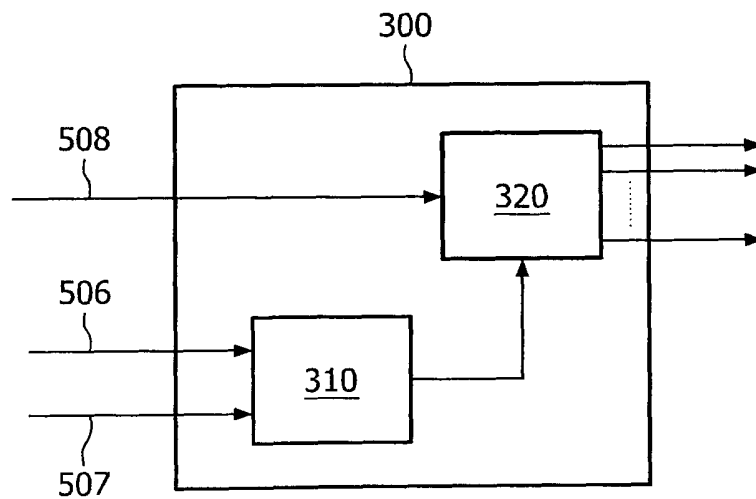


FIG. 5

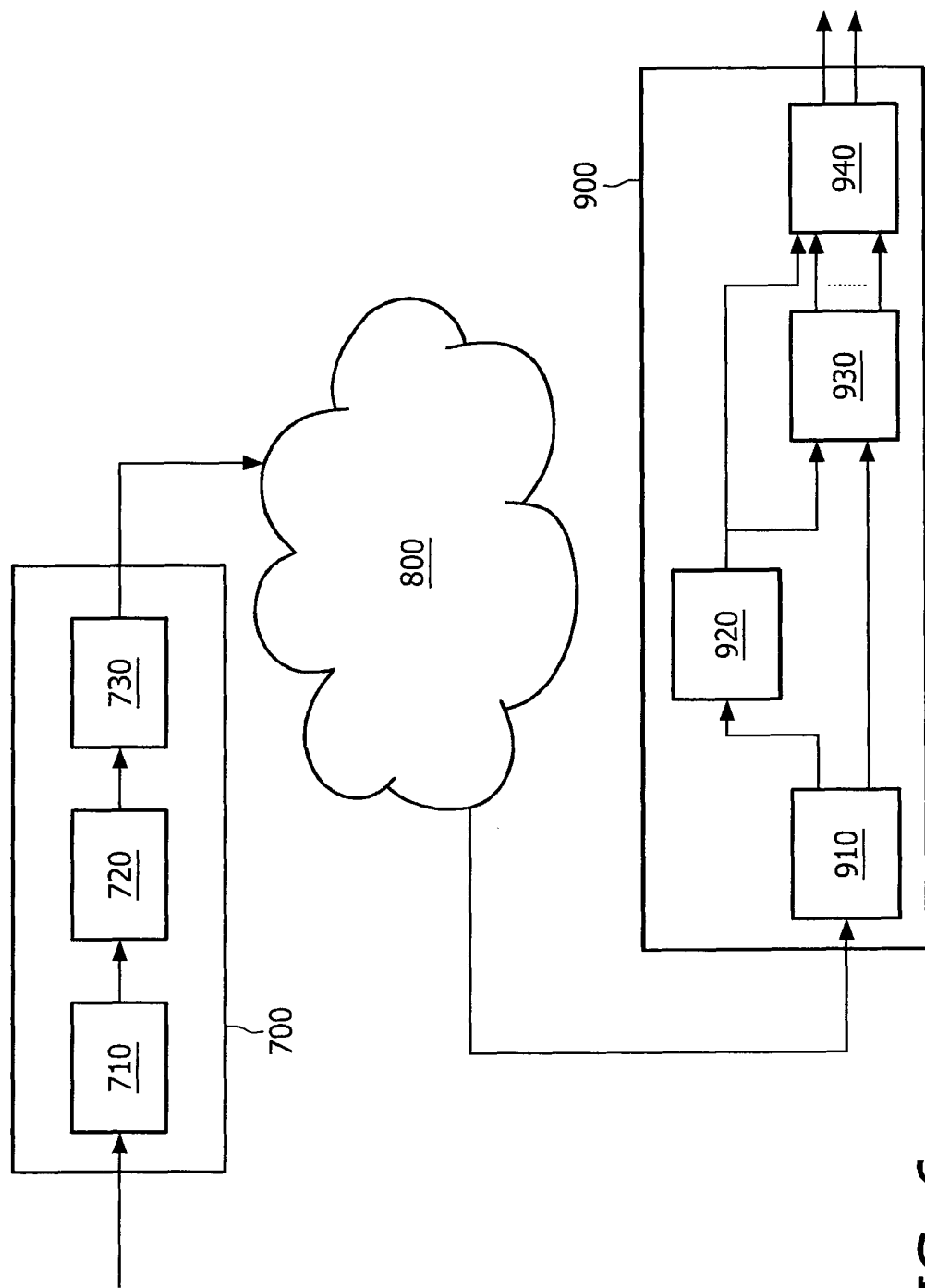


FIG. 6