

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

(21) **PI 0710935-0 A2**

(22) Data de Depósito: 24/11/2007  
(43) Data da Publicação: 14/02/2012  
(RPI 2145)



(51) *Int.Cl.:*  
G10L 19/00

**(54) Título:** MÉTODO PARA CODIFICAR E DECODIFICAÇÃO DE SINAL DE ÁUDIO ORIENTADO A OBJETO E APARELHAGEM PARA O MESMO

**(30) Prioridade Unionista:** 24/11/2006 US 60/860.823, 16/02/2007 US 60/901.642, 22/10/2007 US 60/981.517, 24/10/2007 US 60/982.408, 24/11/2006 US 60/860.823, 16/02/2007 US 60/901.642, 16/02/2007 US 60/901.642, 24/11/2006 US 60/860.823

**(73) Titular(es):** LG Electronics INC

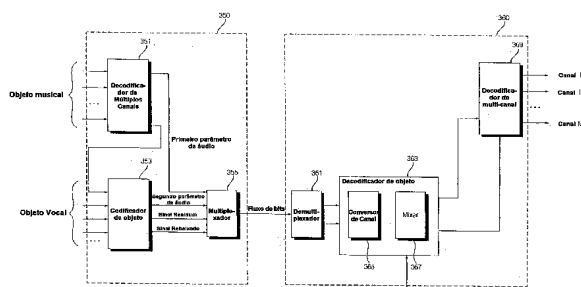
**(72) Inventor(es):** Dong Soo Kim, Hee Suk Pang, Hyun Kook Lee, Jae Hyun Lim, Sung Yong Yoon

**(74) Procurador(es):** Nellie Anne Daniel Shores

**(86) Pedido Internacional:** PCT KR2007005968 de 24/11/2007

**(87) Publicação Internacional:** WO 2008/063034de 29/05/2008

**(57) Resumo:** MÉTODO PARA CODIFICAÇÃO E DECODIFICAÇÃO DE SINAL DE ÁUDIO ORIENTADO A OBJETO E APARELHAGEM PARA O MESMO. A presente invenção refere-se a um método e aparelhagem para a codificação e decodificação de sinais de áudio orientados a objeto. Este método de decodificação de áudio inclui a extração de um primeiro sinal de áudio e de um primeiro parâmetro de áudio aonde um objeto musical é codificado em uma base de canal, e um segundo sinal de áudio e um segundo parâmetro de áudio aonde um objeto vocal é codificado em uma base objeto, a partir de um sinal de áudio, gerando um terceiro sinal de áudio através do emprego de, pelo menos, um dos primeiros e segundos sinais de áudio, e gerando um sinal de áudio de múltiplos canais através do emprego de, pelo menos, um dos primeiros e segundos parâmetros de áudio e o terceiro sinal de áudio. Resultando que, pode-se reduzir de forma eficiente a quantidade de cálculos nos processos de codificação e decodificação e no tamanho de um fluxo de bits que é codificado.



## “MÉTODO PARA CODIFICAÇÃO E DECODIFICAÇÃO DE SINAL DE ÁUDIO ORIENTADO A OBJETO E APARELHAGEM PARA O MESMO”

### Campo Técnico

5 A presente invenção refere-se a um método de codificação e decodificação e aparelhagem para codificação e decodificação de sinais de áudio orientado a objeto de modo que os sinais de áudio possam ser processados através de um agrupamento eficiente.

### Estado da Técnica

10 Em regra, um código de áudio orientado a objeto faz uso de um método de envio da soma de um parâmetro específico extraída a partir de cada sinal de objeto e dos sinais de objeto, restabelecendo os respectivos sinais de objeto a partir do mesmo, e mixando os sinais de objeto tanto quanto seja a quantidade de canais desejada. Assim, quando o número de sinais de objeto se multiplica, a quantidade de informação necessária para se mixar os respectivos sinais de objeto é aumentada na proporção desse número de sinais de objeto.

15 Entretanto, para os sinais de objeto apresentando uma relação funcional próxima, envia-se similar informação de mixagem e, assim por diante, com referência a cada sinal de objeto. Resultando que, a eficiência pode ser melhorada caso os sinais de objeto sejam enfeixados em um grupo e idêntica informação seja enviada somente uma única vez.

20 Mesmo em um método genérico de codificação e decodificação, um efeito semelhante pode ser obtido através do enfeixamento de vários sinais de objeto em um sinal de objeto. Contudo, caso este método seja empregado, a unidade do sinal de objeto é aumentada e também é impossível se misturar o sinal de objeto na forma de uma unidade de sinal de objeto original antes do enfeixamento.

### Descrição da Invenção

#### 25 Problema Técnico

Resulta que, um objetivo da presente invenção é o fornecimento de um método de codificação e decodificação de áudio para a codificação e decodificação de sinais de objeto, aonde os sinais de áudio de objeto com uma associação são enfeixados em um grupo e podem ser processados dessa forma baseados por grupo, e uma aparelhagem para o mesmo.

#### 30 Solução Técnica

Para se chegar ao objetivo acima, um método de decodificação de sinal de áudio de acordo com a presente invenção inclui as etapas de extrair-se um primeiro sinal de áudio e um primeiro parâmetro de áudio aonde um objeto musical é codificado em uma base de canal e, extrair-se um segundo sinal de áudio e um segundo parâmetro de áudio aonde um objeto vocal é codificado em uma base objeto, a partir de um sinal de áudio; geração de um terceiro sinal de áudio através do emprego de, pelo menos, um dos primeiros e segundos

sinais de áudio, e geração de um sinal de áudio em múltiplos canais pelo emprego de pelo menos um dos primeiro e segundo parâmetros de áudio e o terceiro sinal de áudio.

Ainda, para se chegar ao objetivo acima, um método de decodificação de áudio de acordo com a presente invenção inclui as etapas de receber um sinal de mistura descendente, extrair um primeiro sinal de áudio aonde um objeto musical incluindo um objeto vocal apresente-se codificado e um segundo sinal de áudio aonde um objeto vocal apresente-se codificado, a partir do sinal de mistura descendente, e gerar qualquer sinal de áudio incluindo somente o objeto vocal, um sinal de áudio compreendendo o objeto vocal, e um sinal de áudio não incluindo o objeto vocal com base no primeiro e segundo sinais de áudio.

Enquanto isso, uma aparelhagem de decodificação de sinal de áudio de acordo com a presente invenção inclui um multiplexador para extração de um sinal de mistura descendente e informação paralela a partir de um fluxo de bits recebido, um decodificador de objeto para geração de um terceiro sinal de áudio através do emprego de, pelo menos, um primeiro sinal de áudio aonde um objeto musical extraído do sinal de mistura descendente é codificado em uma base de canal e um segundo sinal de áudio aonde um objeto vocal extraído do sinal de mistura descendente é codificado em uma base objeto, e um decodificador de múltiplos canais para geração de um sinal de áudio em múltiplos canais através do emprego de, pelo menos, um primeiro parâmetro de áudio e um segundo parâmetro de áudio, extraídos da informação paralela, e o terceiro sinal de áudio.

Ainda, uma aparelhagem para decodificação de áudio de acordo com a presente invenção que inclui um decodificador de objeto para geração de qualquer sinal de áudio incluindo somente um objeto vocal, um sinal de áudio compreendendo o objeto vocal, e um sinal de áudio não incluindo o objeto vocal orientado em um primeiro sinal de áudio aonde um objeto musical extraído de um sinal de mistura descendente é codificado e um segundo sinal de áudio aonde um objeto vocal extraído do sinal de mistura descendente é codificado, e um decodificador para múltiplos canais para gerar um sinal de áudio em múltiplos canais através do emprego de uma saída de sinal a partir do decodificador de objeto.

Ainda, um método de codificação de áudio de acordo com a presente invenção que inclui as etapas de geração de um primeiro sinal de áudio aonde um objeto musical é codificado em uma base de canal, e um primeiro parâmetro de áudio correspondendo ao objeto musical, gerando um segundo sinal de áudio aonde um objeto vocal é codificado em uma base objeto, e um segundo parâmetro de áudio correspondendo ao objeto vocal, e gerando um fluxo de bits incluindo o primeiro e segundo sinais de áudio, e primeiro e segundo parâmetros de áudio.

De acordo com a presente invenção, é fornecida uma aparelhagem para codificação de áudio incluindo um codificador de múltiplos canais para geração de um

primeiro sinal de áudio aonde um objeto musical é codificado em uma base de canal, e um primeiro parâmetro de áudio orientado a canal com respeito ao objeto musical, um codificador de objeto para geração de um segundo sinal de áudio aonde um objeto vocal é codificado em uma base objeto, e um segundo parâmetro de áudio orientado a objeto com respeito ao objeto vocal, e um multiplexador para geração de um fluxo de bits, incluindo o primeiro e segundo sinais de áudio, e o primeiro e segundo parâmetros de áudio.

Para se chegar ao objetivo acima, a presente invenção fornece uma mídia de gravação fixa aonde é gravado um programa para a execução do método acima em um computador.

#### Efeitos Vantajosos

De acordo com a presente invenção, os sinais de áudio de objeto com uma associação podem ser processados em uma base por grupo, enquanto que fazendo uso das vantagens da codificação e decodificação de sinais de áudio orientados a objeto da melhor maneira possível. Resultando que, pode-se melhorar a eficiência em termos da quantidade de cálculo para os processos de codificação e decodificação, o tamanho de um fluxo de bits que venha a ser codificado, e assim por diante. Ainda, a presente invenção pode ser aplicada num sistema karaokê, etc., plenamente utilizável através do agrupamento dos sinais de objeto em um objeto musical, um objeto vocal, etc.

#### Breve Descrição dos Desenhos

A Figura 1 compreende de um diagrama de blocos de uma aparelhagem de codificação e decodificação de áudio de acordo com uma primeira modalidade da presente invenção;

a Figura 2 compreende de um diagrama de blocos de uma aparelhagem de codificação e decodificação de áudio de acordo com uma segunda modalidade da presente invenção;

a Figura 3 compreende de uma vista ilustrando uma correlação entre uma fonte sonora, grupos, e sinais de áudio;

a Figura 4 compreende de um diagrama de blocos de uma aparelhagem de codificação e decodificação de áudio de acordo com uma terceira modalidade da presente invenção;

as Figuras 5 e 6 compreendem vistas ilustrando um objeto principal e um objeto de fundo;

as Figuras 7 e 8 compreendem vistas ilustrando uma configuração de um fluxo de bits gerado na aparelhagem de codificação;

a Figura 9 compreende de um diagrama de blocos de uma aparelhagem de codificação e decodificação de áudio de acordo com uma quarta modalidade da presente invenção;

a Figura 10 compreende de uma vista ilustrando um caso aonde uma pluralidade de objetos principais é utilizada;

a Figura 11 compreende de um diagrama de blocos de uma aparelhagem de codificação e decodificação de áudio de acordo com uma quinta modalidade da presente invenção;

a Figura 12 compreende de um diagrama de blocos de uma aparelhagem de codificação e decodificação de áudio de acordo com uma sexta modalidade da presente invenção;

a Figura 13 compreende de um diagrama de blocos de uma aparelhagem de codificação e decodificação de áudio de acordo com uma sétima modalidade da presente invenção;

a Figura 14 compreende de um diagrama de blocos de uma aparelhagem de codificação e decodificação de áudio de acordo com uma oitava modalidade da presente invenção;

a Figura 15 compreende de um diagrama de blocos de uma aparelhagem de codificação e decodificação de áudio de acordo com uma nona modalidade da presente invenção; e

a Figura 16 compreende de uma vista ilustrando o caso aonde os objetos vocais são codificados etapa por etapa.

#### Melhor Forma de Condução da Invenção

A presente invenção será agora descrita em detalhes com referência aos desenhos de acompanhamento.

A Figura 1 compreende de um diagrama de blocos de uma aparelhagem de codificação e decodificação de áudio de acordo com uma primeira modalidade da presente invenção. A aparelhagem de codificação e decodificação de áudio de acordo com a presente modalidade decodifica e codifica um sinal de objeto correspondendo a um sinal de áudio orientado aobjeto com base em um conceito de agrupamento. Em outras palavras, os processos de codificação e decodificação são realizados numa base grupal através da conexão de um ou mais sinais de objeto com uma associação no mesmo grupo.

Com referência à Figura 1, apresenta-se uma aparelhagem de codificação de áudio 110 incluindo um codificador de objeto 111, e uma aparelhagem de decodificação de áudio 120 incluindo um decodificador de objeto 121 e um mixer/componente de renderização 123. Embora não mostrada nos desenhos, a aparelhagem de codificação 110 pode incluir um multiplexador, etc. para a geração de um fluxo de bits aonde um sinal de mistura descendente e a informação paralela são combinados, e a aparelhagem de decodificação 120 pode incluir um demultiplexador, etc. para extração de um sinal de mistura descendente e a informação paralela a partir de um fluxo de bits recebido. Esta montagem compreende o

caso com a aparelhagem de codificação e decodificação de acordo com outras modalidades que serão posteriormente descritas.

A aparelhagem de codificação 110 recebe N sinais de objeto, e a informação de grupo, incluindo a informação da posição relativa, informação de tamanho, informação de tempo de retardo, etc. em uma base por grupo, do sinal de objeto com uma associação. A aparelhagem de codificação 110 codifica um sinal, aonde sinais de objeto com uma associação são agrupados, e gera um sinal de mistura descendente orientado a objeto apresentando um ou mais canais e informação paralela, incluído informação extraída de cada sinal de objeto, etc.

Na aparelhagem de decodificação 120, o decodificador de objeto 121 gera sinais, que serão codificados com base no agrupamento, baseado no sinal de mistura descendente e na informação paralela, e o mixer/componente de renderização 123 posiciona os sinais liberados do decodificador de objeto 121 em posições específicas em um espaço de múltiplos canais em um nível específico com base na informação de controle. Ou seja, a aparelhagem de decodificação 120 gera sinais de múltiplos canais sem compactar os sinais, que são codificados com base no agrupamento, em uma base objeto.

Através dessa construção, a quantidade de informação a ser transmitida pode ser reduzida pelo agrupamento e codificação dos sinais de objeto apresentando similares alterações de posição, alterações de tamanho, alteração de retardo, etc. de acordo com o tempo. Ainda, caso os sinais de objeto sejam agrupados, a informação paralela em comum com respeito a um grupo pode ser transmitida, assim pode-se controlar facilmente vários sinais de objeto pertencendo ao mesmo grupo.

A Figura 2 compreende um diagrama de blocos de uma aparelhagem de codificação e decodificação de áudio de acordo com uma segunda modalidade da presente invenção. Uma aparelhagem de decodificação de sinal de áudio 140 de acordo com a presente modalidade é diferente da primeira modalidade no sentido que ela inclui ainda um extrator de objeto 143.

Em outras palavras, a aparelhagem de codificação 130, o decodificador de objeto 141, e o mixer/componente de renderização 145 apresentam a mesma função e construção como aqueles da primeira modalidade. Entretanto, uma vez que a aparelhagem de decodificação 140 inclui ainda o extrator de objeto 143, um grupo, ao qual um sinal de objeto correspondente pertença, pode ser descompactado por base de objeto, quando a descompactação de uma unidade objeto fizer-se necessária. Neste caso, os grupos inteiros não são descompactados em uma base objeto, mas os sinais de objeto podem ser extraídos com respeito somente a grupos aonde não se possa realizar a mixagem de cada grupo, etc.

A Figura 3 compreende de uma vista ilustrando uma correlação entre uma fonte sonora, grupos, e sinais de objeto. Conforme mostrado na Figura 3, sinais de objeto

apresentando uma propriedade semelhante são agrupados de modo que o tamanho de um fluxo de bits possa ser reduzido e todos os sinais de objeto façam parte um grupo superior.

A Figura 4 compreende de um diagrama de blocos de uma aparelhagem de codificação e decodificação de acordo com uma terceira modalidade da presente invenção.

5 Na aparelhagem de codificação e decodificação de áudio de acordo com a modalidade presente, utiliza-se o conceito de um canal central para mixagem de mistura descendente.

Com referência a Figura 4, apresenta-se um codificador de objeto 151 pertencendo a uma aparelhagem de codificação de áudio, e uma aparelhagem de decodificação de áudio 160 incluindo um decodificador de objeto 161 e um mixer/componente de renderização 163.

10 O codificador de objeto 151 recebe N sinais de objeto ( $N > 1$ ) e gera sinais que são de mistura descendente nos canais M ( $1 < M < N$ ). Na aparelhagem de decodificação 160, o decodificador de objeto 161 decodifica os sinais, que são então de mistura descendente nos canais M, e daí para os sinais de objeto N novamente, e o mixer/componente de renderização 163, finalmente, libera os sinais do canal L ( $L \geq 1$ ).

15 A esta altura, os canais de mistura descendente M gerados pelo codificador de objeto 151 compreendem de K canais centrais de mistura descendente ( $K < M$ ) e os canais M-K não-principais de mistura descendente. A razão do porque dos canais de mistura descendente serem construídos conforme a descrição anterior é que a importância dos mesmos pode ser alterada de acordo com um sinal de objeto. Em outras palavras, um  
20 método genérico de codificação e decodificação não apresenta uma resolução suficiente com respeito a um sinal de objeto e, portanto, inclui os componentes de outros sinais de objeto numa base por sinal de objeto. Assim, caso os canais de mistura descendente compreendam os canais principais de mistura descendente e os canais não principais de mistura descendente, conforme descrito acima, pode-se minimizar a interferência entre os  
25 sinais de objeto.

Nesta situação, o canal principal de mistura descendente pode utilizar um método de processamento diferente daquele para o canal não-principal de mistura descendente. Por exemplo, na Figura 4, a entrada de informação paralela junto ao mixer/componente de renderização 163 pode ser definida somente no canal principal de mistura descendente. Em  
30 outras palavras, o mixer/componente de renderização 163 pode ser configurado para controlar somente sinais de objeto decodificados do canal principal de mistura descendente e não os sinais de objeto decodificados a partir do canal não-principal de mistura descendente.

Como outro exemplo, pode-se construir o canal principal de mistura descendente  
35 somente com um pequeno número de sinais de objeto, e os sinais de objeto podem ser agrupados e daí, controlados com base em uma informação de controle. Por exemplo, um canal principal de mistura descendente a mais pode ser construído somente com sinais

vocais de maneira a se montar um sistema karaokê. Ainda, um canal principal de mistura descendente adicional pode ser construído através do agrupamento somente de sinais de uma bateria, etc. de modo que a intensidade de um sinal de baixa frequência, tal como um sinal de bateria, possa ser controlado com precisão.

5 Enquanto isso, a música é gerada, em regra, através da mixagem de vários canais de áudio apresentando o formato de uma trilha, etc. Por exemplo, no caso de música apresentando bateria, guitarra, piano, e sinais de vocais, cada um dos sinais de bateria, guitarra, piano e vocal pode se tornar um sinal de objeto. Neste caso, um sinal, do total de  
10 sinais de objeto, que seja determinado como sendo especificamente importante e que possa ser controlado por um usuário, ou um número de sinais de objeto, que sejam mixados e controlados como aquele sinal de objeto, pode ser definido com um objeto principal. Ainda, a mixagem de sinais de objeto que não compreenda aquele objeto principal do total de sinais de objeto pode ser definida como sendo um objeto de fundo. De acordo com esta definição, pode-se dizer que um objeto total ou um objeto musical consiste do objeto principal e do  
15 objeto de fundo.

As Figuras 5 e 6 compreendem vistas ilustrando o objeto principal e o objeto de fundo. Conforme mostrado na Figura 5a, assumindo que o objeto principal seja um som vocal e o objeto de fundo compreenda a mixagem de sons de todos os instrumentos musicais que não compreendam o som vocal, um objeto musical pode incluir um objeto  
20 vocal e um objeto de fundo do som mixado dos instrumentos musicais que não inclua o som vocal. A quantidade de objetos principais pode variar de um ou mais, conforme mostrado na Figura 5b.

Ainda, o objeto principal pode apresentar um formato em que vários sinais de objeto são mixados. Por exemplo, conforme mostrado na Figura 6, a mixagem de som de vocal e  
25 guitarra pode ser utilizada como os objetos principais e a sonorização dos instrumentos musicais restantes pode ser utilizada como objetos de fundo.

De forma a se controlar separadamente o objeto principal e o objeto de fundo no objeto musical, o fluxo de bits codificado na aparelhagem de codificação pode apresentar um dos formatos mostrados na Figura 7.

30 A Figura 7a ilustra uma situação aonde o fluxo de bits gerado na aparelhagem de codificação compreende de um fluxo de bits musical e um fluxo de bits de objeto principal. O fluxo de bits musical apresenta um formato aonde todos os sinais de objeto são mixados, e refere-se a um fluxo de bits correspondendo a soma de todos os objetos principais e objetos de fundo. A Figura 7b ilustra uma situação aonde o fluxo de bits compreende de um fluxo de  
35 bits musical e um fluxo de bits de objeto de fundo. A Figura 7c ilustra uma situação aonde o fluxo de bits compreende de um fluxo de bits de objeto principal e um fluxo de bits de objeto de fundo.

Na Figura 7, é concebida uma regra no sentido de se gerar o fluxo de bits musical, o fluxo de bits do objeto principal, e o fluxo de bits do objeto de fundo utilizando-se um codificador e decodificador apresentando o mesmo método. Contudo, quando se é utilizado o objeto principal na forma de objeto vocal, o fluxo de bits musical pode ser codificado e decodificado utilizando-se o MP3, e o fluxo de bits de objeto vocal pode ser decodificado e codificado utilizando-se um codificador de voz, tal como o AMR, QCELP, EFR, ou o EVRC de maneira a se reduzir a capacidade do fluxo de bits. Em outras palavras, os métodos de codificação e decodificação do objeto musical e do objeto principal, do objeto principal e do objeto de fundo, e assim por diante, podem diferir.

Na Figura 7a, a parte do fluxo de bits musical é configurada utilizando-se o mesmo método como o método de codificação genérico. Ainda, no método de codificação genérico, tal como o MP3 ou AAC, uma parte – aonde informação paralela é incluída, tal como uma região autônoma ou uma região auxiliar – é incluída na metade final do fluxo de bits. O fluxo de bits do objeto principal pode ser adicionado nesta parte. Entretanto, um fluxo de bits completo pode compreender de uma região aonde o objeto musical apresente-se codificado, e uma região de objeto principal posterior a região aonde o objeto musical apresente-se codificado. Nesta altura, introduz-se um indicador, sinalizador ou algo semelhante, informando que foi adicionado o objeto principal, podendo este ser adicionado junto à primeira metade da região paralela, de modo que, caso exista o objeto principal na aparelhagem de decodificação, ele possa ser determinado.

O caso da Figura 7b compreende basicamente do mesmo formato da Figura 7a. Na Figura 7b, utiliza-se o objeto de fundo ao invés do objeto principal da Figura 7a.

A Figura 7c ilustra um caso aonde o fluxo de bits compreende de um fluxo de bits de objeto principal e um fluxo de bits de objeto de fundo. Nesta situação, o objeto musical compreende da soma ou mixagem do objeto principal e do objeto de fundo. Em um método de configuração do fluxo de bits, o objeto de fundo pode ser armazenado primeiramente, e o objeto principal pode ser armazenado, então, na região auxiliar. Alternativamente, o objeto principal pode ser armazenado primeiramente e o objeto de fundo pode então ser armazenado na região auxiliar. Em tal situação, um indicador para dar a informação sobre a região paralela pode ser adicionado junto à primeira metade da região paralela, que é idêntica a descrita anteriormente.

A Figura 8 ilustra um método de configuração de fluxo de bits de modo que o objeto principal que tenha sido adicionado possa ser determinado. Um primeiro exemplo é aquele aonde, após a conclusão de um fluxo de bits musical, a região correspondente compreende de uma região auxiliar até que tenha início um quadro seguinte. No primeiro exemplo, pode ser incluído somente um indicador, informando que objeto principal tenha sido codificado.

Um segundo exemplo corresponde a um método de codificação solicitando um indicador, informando que tem início uma região auxiliar ou região de dados após a conclusão do fluxo de bits musical. Para esta finalidade, na codificação de um objeto principal, são requeridos dois tipos de indicadores, tal como um indicador para informar o início da região auxiliar e um indicador para informar o objeto principal. Na decodificação deste fluxo de bits, o tipo de dado é determinado pela leitura do indicador, e o decodificador é então decodificado através da leitura de um dado a parte.

Figura 9 compreende de um diagrama de blocos de uma aparelhagem de codificação e decodificação de áudio, de acordo com uma quarta modalidade da presente invenção. A aparelhagem de codificação e decodificação de áudio, de acordo com a presente modalidade codifica e decodifica um fluxo de bits aonde se é adicionado um objeto vocal na forma de um objeto principal.

Com referência a Figura 9, um codificador 211 incluído em uma aparelhagem de codificação codifica um sinal musical, incluindo um objeto vocal e um objeto musical. Exemplos de sinais musicais do codificador 211 podem incluir o MP3, AAC, WMA, e assim por diante. O codificador 211 adiciona o objeto vocal junto ao fluxo de bits na forma de um objeto principal, que não compreenda quaisquer sinais musicais. Nesta altura, o codificador 211 adiciona o objeto vocal junto a uma parte, alimentado a informação paralela, tal como uma região autônoma ou uma região auxiliar, conforme mencionado anteriormente, e adiciona também um indicador, etc., informando à aparelhagem de codificação do fato de que existe, adicionalmente, o objeto vocal, junto à parte.

Uma aparelhagem de decodificação 220 inclui um decodificador de código genérico 221, um decodificador vocal 223, e um mixer 225. O decodificador de código genérico 221 decodifica a parte de fluxo de bits musical do fluxo de bits recebido. Nesta situação, uma região de objeto principal é, simplesmente, reconhecida na forma de uma região paralela ou uma região de dados, mas não é utilizada no processo de decodificação. O decodificador vocal 223 decodifica a parte do objeto vocal do fluxo de bits recebido. O mixer 225 mixa os sinais decodificados no decodificador de código genérico 221 e no decodificador vocal 223 e libera o resultado da mixagem.

Quando é recebido um fluxo de bits, aonde um objeto vocal esteja incluído na forma de um objeto principal, a aparelhagem de codificação, não incluído o decodificador vocal 223, decodifica somente um fluxo de bits musical e libera os resultados da decodificação. Entretanto, mesmo nessa situação, é a mesma coisa como no caso para a liberação do áudio em geral, uma vez que o sinal de vocal é incluído no fluxo musical.

Ainda, no processo de decodificação, determina-se se o objeto vocal foi adicionado ao fluxo de bits com base em um indicador, etc. Quando se é impossível se decodificar o

objeto vocal, esse objeto vocal é descartado através de salto, etc., mas quando é possível se decodificar o objeto vocal, o objeto vocal é decodificado e empregado para mixagem.

O decodificador de código genérico 221 é adaptado para tocar música e, em regra, utiliza decodificação de áudio. Por exemplo, existe o MP3, o AAC, o HE-AAC, WMA, o Ogg Vorbis, e outros. O decodificador vocal 223 pode fazer uso do mesmo código, ou ser diferente, daquele empregado pelo decodificador de código genérico 221. Por exemplo, o decodificador vocal 223 pode fazer uso de um código de voz, tal como o EVRC, EFR, AMR ou o QCELP. Neste caso, a quantidade de cálculos para a decodificação pode ser reduzida.

Ainda, caso o objeto vocal seja em mono, a taxa de bits pode ser reduzida para a maior extensão possível. Entretanto, caso o fluxo de bits musical não possa ser CARACTERIZADO somente como mono devido a apresentar-se em canais estéreos e os sinais vocais nos canais à esquerda e direita diferirem, o objeto vocal pode caracterizar-se também por uma saída em estéreo.

Na aparelhagem de decodificação 220, de acordo com a modalidade presente, reproduz-se qualquer um dos modos aonde somente música seja tocada, um modo em que seja reproduzido somente o objeto principal, e um modo, aonde a música e um objeto musical são mixados e tocados de forma adequada, podendo ser selecionados e reproduzidos em resposta a um comando de controle de usuário, tal como uma tecla ou um menu de manipulação em um dispositivo reprodutor.

Na eventualidade de que seja descartado um objeto principal e somente seja tocada música original, isto corresponde a reproduzir-se a música existente. Contudo, uma vez que é possível se mixar em resposta a um comando de controle de usuário, etc., o tamanho do objeto principal ou de um objeto de fundo, etc. pode ser controlado. Quando o objeto principal representa um objeto vocal, isto significa que somente o vocal pode ser aumentado ou diminuído em comparação com o fundo musical.

Um exemplo de quando somente um objeto musical é tocado pode incluir um em que um objeto vocal ou um som especial de instrumento musical é utilizado como o objeto principal. Em outras palavras, isto significa que somente é ouvido o vocal sem a música de fundo, somente o som do instrumento musical é ouvido sem o fundo musical, e assim por diante.

Quando a música e um objeto principal são mixados e ouvidos de forma adequada, isto significa que somente o vocal é aumentado ou diminuído quando comparado ao fundo musical. Em particular, na eventualidade que os componentes vocais sejam completamente eliminados da música, a música pode ser usada como um sistema karaokê, uma vez que os componentes vocais desaparecem. Caso se codifique um objeto vocal na aparelhagem de codificação em um estado em que seja revertida a fase do objeto vocal, a aparelhagem de

decodificação pode reproduzir um sistema de karaokê através da adição do objeto vocal a um objeto musical.

No processo acima, descreveu-se que o objeto musical e o objeto principal são, respectivamente, decodificados, e daí, mixados. Contudo, o processo de mixagem pode ser realizado durante o processo de decodificação. Por exemplo, na série de transformada de codificação, tal como uma MDCT (Transformada Modificada Discreta de Coseno) incluindo MP3 e AAC, a mixagem pode ser realizada nos coeficientes DCT e a inversa MDCT pode ser, finalmente, efetuada, gerando dessa maneira as saídas de PCM. Neste caso, a quantidade total de cálculos pode ser reduzida de forma significativa. Ainda, a presente invenção não está limitada ao MDCT, mas inclui todas as transformadas aonde os coeficientes são mixados em um domínio de transformadas com respeito a um decodificador para séries de codificação de transformadas em geral, sendo então a decodificação realizada.

Mais ainda, um exemplo em que é empregado um objeto principal foi descrito no exemplo anterior. Entretanto, um número de objetos principais pode ser utilizado. Por exemplo, conforme mostrado na Figura 10, o vocal pode ser utilizado na forma de um objeto principal 1 e uma guitarra pode ser utilizada na forma de um objeto principal 2. Esta construção é bastante útil quando somente um objeto de fundo que não seja o vocal e uma guitarra esteja sendo tocado na música e um usuário pode diretamente desempenhar o vocal e uma guitarra. Ainda, os fluxos de bits podem ser reproduzidos através de várias combinações musicais, um das quais o vocal pode ser retirado da música, e numa outra, uma guitarra pode ser retirada da música, numa outra o vocal e a guitarra são retirados da música, e assim por diante.

Enquanto isso, na presente invenção, um canal indicado por um fluxo de bits de vocal pode ser expandido. Por exemplo, partes inteiras da música, uma parte do som da bateria, ou uma parte aonde somente o som da bateria é retirado das outras partes integrais na música podem ser reproduzidas utilizando-se somente um fluxo de bits de bateria. Ainda, a mixagem pode ser controlada em uma base por partes utilizando-se dois ou mais fluxos de bits adicionais, tal como o fluxo de bits de vocal e o fluxo de bits de bateria.

Ainda, na presente modalidade, somente descreveu-se a condição estéreo/mono. Contudo, a presente modalidade pode ser expandida para uma situação de múltiplos canais. Por exemplo, um fluxo de bits pode ser configurado através da adição de um objeto vocal, um fluxo de bits de objeto principal, e assim por diante junto a um fluxo de bits de canal 5,1, e mediante reprodução, pode ser tocado qualquer som original, som de onde o vocal foi retirado, e som que inclua somente vocal.

A presente modalidade pode ser também configurada para sustentar somente música e um modo aonde o vocal seja retirado da música, mas não para sustentar um modo

aonde somente vocal (um objeto principal) seja reproduzido. Este método pode ser utilizado quando cantores não desejam que somente o vocal seja reproduzido. Ele pode ser expandido para a configuração de um decodificador aonde um identificador, indicando se uma função que sustenta somente vocal existe ou não, é posicionado em um fluxo de bits e a faixa de reprodução é decidida com base no fluxo de bits.

A Figura 11 compreende de um diagrama de blocos de uma aparelhagem de codificação e decodificação de áudio de acordo com uma quinta modalidade da presente invenção. A aparelhagem de codificação e decodificação de áudio de acordo com a modalidade presente pode implementar um sistema de karaokê utilizando um sinal residual.

Quando especializando um sistema de karaokê, um objeto musical pode ser dividido em um objeto de fundo e um objeto principal, conforme mencionado anteriormente. O objeto principal refere-se a um sinal de objeto que será controlado separadamente do objeto de fundo. Em particular, o objeto principal pode se referir a um sinal de objeto vocal. O objeto de fundo compreende a soma dos sinais integrais de objeto, excluído o objeto principal.

Com referência a Figura 11, um codificador 251 incluído em uma aparelhagem de codificação codifica um objeto de fundo e um objeto principal, com estes sendo colocados conjuntamente. No instante da codificação, pode ser utilizado um código de áudio genérico, tal como o AAC ou o MP3. Caso o sinal seja decodificado em uma aparelhagem de decodificação 260, o sinal decodificado inclui tanto um sinal de objeto de fundo quanto um sinal de objeto principal. Assumindo-se que o sinal decodificado seja um sinal decodificado original, o método a seguir pode ser utilizado de forma a aplicar um sistema karaokê junto ao sinal.

O objeto principal é incluído em um fluxo de bits total na forma de um sinal residual. O objeto principal é decodificado e daí subtraído do sinal de decodificação original. Neste caso, um primeiro decodificador 261 decodifica o sinal total e o segundo decodificador 263 decodifica o sinal residual, aonde  $g = 1$ . Alternativamente, o sinal de objeto principal apresentando uma fase reversa pode ser incluído no fluxo de bits total na forma de um sinal residual. O sinal de objeto principal pode ser decodificado e daí, adicionado junto ao sinal de decodificação original. Neste caso,  $g = -1$ . Em qualquer situação, um tipo de sistema karaokê escalável é possível através do controle do valor de  $g$ .

Por exemplo, quando  $g = -0,5$  ou  $g = 0,5$ , o objeto principal ou o objeto vocal não são plenamente removidos, mas somente o nível pode ser controlado. Ainda, caso o valor de  $g$  seja ajustado para um valor positivo ou negativo, ocorre um efeito em que o tamanho do objeto vocal pode ser controlado. Caso o sinal de decodificação original não seja utilizado e somente seja liberado o sinal residual, pode-se sustentar um modo solo somente com vocal.

A Figura 12 compreende de um diagrama de blocos de uma aparelhagem de codificação e decodificação de áudio de acordo com uma sexta modalidade da presente invenção. A aparelhagem de codificação e decodificação de áudio de acordo com a presente modalidade faz uso de dois sinais residuais através da diferenciação dos sinais residuais para uma saída de sinal para karaokê e uma saída em modo vocal.

Com referência a Figura 12, um sinal de decodificação original codificado em um primeiro decodificador 291 é dividido em um sinal de objeto de fundo e um sinal de objeto principal e daí, liberado em uma unidade de separação de objeto 295. Em realidade, o objeto de fundo inclui parte dos componentes do objeto principal, bem como o objeto de fundo original, e o objeto principal inclui ainda alguns componentes do objeto de fundo, bem como o objeto principal original. Isto ocorre devido a não ser completo o processo de divisão do sinal de decodificação original em objeto de fundo e sinal de objeto principal.

Em particular, com respeito ao objeto de fundo, os componentes do objeto principal inclusos no objeto de fundo podem ser previamente incluídos no fluxo de bits total na forma de sinal residual, o fluxo de bits total pode ser decodificado, e os componentes do objeto principal podem ser, então, subtraídos a partir do objeto de fundo. Neste caso, na Figura 12,  $g = 1$ . Alternativamente, uma fase reversa pode ser fornecida junto aos componentes do objeto principal incluídos no objeto de fundo, os componentes de objeto principal podem ser inclusos no fluxo de bits total na forma de um sinal residual, e o fluxo de bits total pode ser decodificado e, daí, adicionado junto ao sinal do objeto de fundo. Neste caso, na Figura 12,  $g = -1$ . Em qualquer situação, um sistema de karaokê escalável é possível através do controle do valor de  $g$ , conforme mencionado acima em conjunto com a quinta modalidade.

Da mesma maneira, um modo solo pode ser sustentado através do controle de um valor para  $g_1$  após a aplicação do sinal residual junto ao sinal de objeto principal. O valor de  $g_1$  pode ser aplicado conforme descrito anteriormente em consideração da comparação de fase do sinal residual e do objeto original e a gradação do modo vocal.

A Figura 13 compreende de um diagrama de blocos de uma aparelhagem de codificação e decodificação de áudio de acordo com uma sétima modalidade da presente invenção. Na presente modalidade, utiliza-se o método a seguir de maneira a reduzir mais ainda a taxa de bits de um sinal residual na modalidade anterior.

Quando um sinal de objeto principal encontra-se em mono, uma unidade de conversão de estéreo para três canais 305 desempenha a transformada de três canais para estéreo em um sinal estéreo original decodificado em um primeiro decodificador 301. Uma vez que a transformada de três canais para estéreo não é completa, um objeto de fundo (ou seja, uma saída do mesmo) inclui alguns componentes do objeto principal, bem como componentes do objeto de fundo, e um objeto principal (ou seja, uma outra saída do

mesmo) inclui também alguns componentes de objeto de fundo, bem como os componentes de objeto principal.

Então, um segundo decodificador 303 desempenha a decodificação (ou após a decodificação, a conversão qmf ou conversão mdct-para-qmf) em uma parte residual do fluxo de bits total e realiza a soma de ponderação do sinal de objeto de fundo e o sinal de objeto principal. Resultando que, podem ser obtidos os sinais que compreendem, respectivamente, os componentes de objeto de fundo e os componentes de objeto principal.

A vantagem deste método é que uma vez que o sinal de objeto de fundo e o sinal do objeto principal tenham sido divididos através da conversão de estéreo para três canais, um sinal residual para a remoção dos outros componentes incluídos no sinal (ou seja, os componentes do objeto principal restando no interior do sinal de objeto de fundo e os componentes do objeto de fundo restando no interior do sinal de objeto principal ) pode ser construído utilizando-se uma taxa de bits menor.

Com referência a Figura 13, assumindo-se que o componente de objeto de fundo seja B e o componente do objeto principal seja m no interior do sinal de objeto de fundo BS, e o componente de objeto principal seja M, e o componente de objeto de fundo seja b no interior do sinal de objeto principal MS, é estabelecida a fórmula a seguir:

Figura Matemática 1

$$BS = B + m$$

$$MS = M + b$$

Por exemplo, quando o sinal residual R é compreendido de b-m, uma saída final de karaokê KO resulta em:

Figura Matemática 2

$$KO = BS + R = B + b$$

Uma saída em modo solo SO resulta em:

Figura Matemática 3

$$SO = BS - R = M + m$$

O sinal do sinal residual pode ser invertido na fórmula anterior, ou seja,  $R = m - b$ ,  $g = -1$  &  $g1 = 1$ .

Quando configurando-se o BS e o MS, os valores de g e g1 podem ser calculados facilmente, aonde os valores finais de KO e SO irão compreender de B e b, e M e m, dependendo de como os sinais de B, m, M, e/ou b sejam fixados. Nos casos acima, tanto os sinais de karaokê como solo são ligeiramente alterados dos sinais originais, mas saídas de sinal com alta qualidade, que venham a ser presentemente utilizadas, são possíveis devido a que a saída de karaokê não inclui os componentes solos, e a saída solo não inclui os componentes de karaokê.

Ainda, quando dois ou mais objetos principais existem, a conversão de dois para três canais e, uma incrementação/redução do sinal residual, podem ser utilizadas, etapa por etapa.

A Figura 14 compreende de um diagrama de blocos de uma aparelhagem de codificação e decodificação de acordo com uma oitava modalidade da presente invenção. Uma aparelhagem de decodificação de sinal de áudio 290 de acordo com a modalidade presente apresenta-se diferente da sétima modalidade em que se executa duas vezes uma conversão de mono-para-estéreo em cada canal estéreo original, quando um sinal de objeto principal compreende de um sinal estéreo.

Uma vez que a conversão de mono para estéreo não é perfeita também, um sinal de objeto de fundo (ou seja, uma saída do mesmo) inclui parte dos componentes de objeto principal, bem como partes dos componentes de objeto de fundo, e um sinal de objeto principal (ou seja, a outra saída do mesmo) inclui também alguns componentes de objeto de fundo, bem como os componentes de objeto principal. Posteriormente, a decodificação (ou após a decodificação, a conversão qmf ou mdct-para-qmf) é realizada em uma parte residual de um fluxo de bits total, e os seus componentes de canal esquerdo e direito são então adicionados junto aos canais da esquerda e da direita de um sinal de objeto de fundo e um sinal de objeto principal, respectivamente, que são multiplicados por um peso, de modo possam ser obtidos os sinais compreendendo de um componente de objeto de fundo (estéreo) e um componente de objeto principal (estéreo).

Na eventualidade de que os sinais residuais em estéreo sejam formados através do emprego da diferença entre os componentes à esquerda e à direita do objeto de fundo estéreo e o objeto principal estéreo,  $g = g_2 = -1$ , e  $g_1 = g_3 = 1$  na Figura 14. Ainda, conforme descrito acima, os valores de  $g$ ,  $g_1$ ,  $g_2$ , e  $g_3$  podem ser calculados facilmente de acordo com os sinais do sinal de objeto de fundo, o sinal de objeto principal, e o sinal residual.

Em regra, um sinal de objeto principal pode ser mono ou estéreo. Por esta razão, um sinalizador indicando se o sinal do objeto principal é mono ou estéreo, é colocado no interior do fluxo de bits total. Quando o sinal de objeto principal é mono, o sinal de objeto principal pode ser decodificado utilizando-se o método descrito em conjunto com a sétima modalidade da Figura 13, e quando o sinal de objeto principal é estéreo, o sinal de objeto principal pode ser decodificado utilizando-se o método descrito em conjunto com a oitava modalidade da Figura 14, através da leitura do sinalizador.

Mais ainda, quando são inclusos um ou mais objetos principais, os métodos acima podem ser usados consecutivamente dependendo de se cada um dos objetos principais é mono ou estéreo. Nesta oportunidade, o número de vezes em que cada método é utilizado é idêntico ao número de objetos principais em mono/estéreo. Por exemplo, quando o número de objetos principais é 3, o número de objetos principais em mono dos três objetos

principais é 2, e o número de objetos principais em estéreo é 1, os sinais de karaokê podem ser liberados utilizando-se duas vezes o método descrito em conjunção com a sétima modalidade, e utilizando-se uma vez o método descrito em conjunção com a oitava modalidade da Figura 14. Nesta altura, pode ser decidida previamente a sequência do método descrito em conjunção com a sétima modalidade e o método descrito em conjunto com a oitava modalidade. Por exemplo, o método descrito em conjunção com a sétima modalidade pode ser sempre executado em objetos principais em mono, e o método descrito em conjunção com a oitava modalidade pode ser então executado em objetos principais estereofônicos. Na forma de um outro método de decisão em sequência, um componente de descrição, descrevendo a sequência do método descrito em conjunto com a sétima modalidade e o método descrito em conjunto com a oitava modalidade, pode ser posicionado no interior de um fluxo de bits total e os métodos podem ser executados de maneira seletiva com base no componente de descrição.

A Figura 15 compreende de um diagrama de blocos de uma aparelhagem de codificação e decodificação de áudio de acordo com uma nona modalidade da presente invenção. A aparelhagem de codificação e decodificação de áudio de acordo com a presente modalidade gera objetos musicais ou objetos de fundo utilizando os codificadores de múltiplos canais.

Observa-se pela Figura 15, aparelhagens de codificação de áudio incluindo um codificador de múltiplos canais 351, um codificador de objeto 353, e um multiplexador 355, e uma aparelhagem de decodificação de áudio 360 incluindo um demultiplexador 361, um decodificador de objeto 363, e um decodificador de múltiplos canais 369. O decodificador de objeto 363 pode incluir um conversor de canal 365 e um mixer 367.

O codificador de múltiplos canais 351 gera um sinal, que é de mistura descendente fazendo-se uso de objetos musicais na forma de uma base de canal, e da informação do primeiro parâmetro de áudio orientado a canal através de extração de informação sobre o objeto musical. O decodificador de objeto 353 gera um sinal de mistura descendente, que é codificado utilizando-se objetos vocais e o sinal de mistura descendente a partir do codificador de múltiplos canais 351, na forma de uma base de objeto, de informação do segundo parâmetro de áudio orientado a objeto, e sinais residuais correspondendo aos objetos vocais. O multiplexador 355 gera um fluxo de bits aonde o sinal de mistura descendente gerado a partir do codificador de objeto 353 e a informação paralela são combinados. Nesta oportunidade, a informação paralela é a informação incluindo o primeiro parâmetro de áudio gerado a partir do codificador de múltiplos canais 351, dos sinais residuais e do segundo parâmetro de áudio gerado pelo decodificador de objeto 353, e assim por diante.

Na aparelhagem de decodificação de áudio 360, o demultiplexador 361 faz a demultiplexação do sinal de mistura descendente e a informação paralela no fluxo de bits recebido. O decodificador de objeto 363 gera sinais de áudio com componentes vocais controlados através da utilização de, pelo menos, um sinal de áudio aonde seja codificado o objeto musical em uma base de canal, e um sinal de áudio aonde seja codificado o objeto vocal. O decodificador de objeto 363 inclui o conversor de canal 365 e, portanto, pode desempenhar a conversão de mono para estéreo ou a conversão de dois para três no processo de decodificação. O mixer 367 pode controlar o nível, a posição, etc. de um sinal de objeto específico utilizando um parâmetro de mixagem, etc., que são incluídos na informação de controle. O decodificador de múltiplos canais 369 gera sinais de múltiplos canais utilizando o sinal de áudio e a informação paralela decodificada no decodificador de objeto 361, e assim por diante.

O decodificador de objeto 363 pode gerar um sinal de áudio correspondendo a qualquer um dos modos de karaokê aonde são gerados os sinais de áudio sem os componentes vocais, um modo solo aonde são gerados os sinais de áudio incluindo somente componentes vocais, e um modo genérico aonde são gerados sinais de áudio incluindo componentes vocais de acordo com a informação de controle de entrada.

A Figura 16 compreende de uma vista ilustrando o caso aonde são codificados os objetos vocais, etapa por etapa. Com referência a Figura 16, uma aparelhagem de codificação 380 de acordo com a presente modalidade inclui um codificador de múltiplos canais 381, do primeiro ao terceiro dos decodificadores de objeto 383, 385, e 387, e um multiplexador 389.

O codificador de múltiplos canais 381 apresenta construção idêntica e funciona conforme aqueles codificadores de múltiplos canais mostrados na Figura 15. Vê-se que a presente modalidade difere da nona modalidade da Figura 15 no sentido de que os três codificadores de objeto 383, 385, e 387 são configurados para agruparem objetos vocais etapa por etapa e os sinais residuais, que são gerados nas respectivas etapas de agrupamento, são incluídos em um fluxo de bits gerados pelo multiplexador 389.

No evento de que o fluxo de bits gerado por este processo seja decodificado, um sinal, com os componentes vocais ou outros componentes vocais desejados controlados, pode ser gerado através da aplicação dos sinais residuais, que são extraídos a partir do fluxo de bits, junto a um sinal de áudio codificado pelo agrupamento de objetos musicais ou um sinal de áudio codificado pelo agrupamento dos objetos vocais, etapa por etapa.

Enquanto isso, na modalidade acima, é efetivada uma posição aonde a soma ou a diferença do sinal de decodificação original e o sinal residual, ou a soma ou a diferença do sinal de objeto de fundo ou o sinal do objeto principal e o sinal residual não se apresentam limitadas a um domínio específico. Por exemplo, este processo pode ser realizado em um

domínio temporal ou para um certo domínio da frequência, tal como um domínio MDCT. Alternativamente, este processo pode ser realizado em um domínio de sub-faixa, tal como, um domínio de sub-faixa QMF ou um domínio de sub-faixa híbrido. Em particular, quando este processo é realizado no domínio da frequência ou no domínio de sub-faixas, pode ser gerado um sinal de karaokê escalável através do controle do número de faixas excluindo-se os componentes residuais. Por exemplo, quando o número de sub-faixas de um sinal de decodificação original é 20, caso o número de faixas de um sinal residual seja estabelecido como 20, pode ser liberado um perfeito sinal de karaokê. Quando somente 10 baixas frequências são abrangidas, os componentes vocais são excluídos somente das partes das baixas frequências, permanecendo as partes das altas frequências. No último caso, a qualidade sonora pode ser mais baixa do que no caso formal, mas ocorre uma vantagem no sentido de que a taxa de bits pode ser abaixada.

Ainda, quando o número de objetos principais difere da unidade, os vários sinais residuais podem ser incluídos em um fluxo de bits total e a soma ou a diferença dos sinais residuais pode ser efetuada várias vezes. Por exemplo, quando dois objetos principais incluem vocal e guitarra e os seus sinais residuais são incluídos em um fluxo de bits total, um sinal de karaokê, a partir de onde, tanto os sinais de guitarra e como os de vocal tenham sido removidos, pode ser gerado de tal maneira que primeiramente o sinal de vocal é removido do sinal total e o sinal de guitarra é então removido. Nesta situação, pode ser gerado um sinal de karaokê, a partir de onde somente o sinal de vocal tenha sido removido e um sinal de karaokê a partir de onde somente o sinal de guitarra tenha sido removido. Alternativamente, somente o sinal vocal pode ser liberado ou somente o sinal de guitarra pode ser liberado.

Ainda, de modo a se gerar o sinal de karaokê através da remoção somente do sinal de vocal a partir, fundamentalmente, do sinal total, o sinal total e o sinal de vocal são codificados, respectivamente. São necessários os dois tipos de seções seguintes, de acordo com o tipo de código utilizado para a codificação. Primeiramente, sempre é empregado o mesmo código de codificação no sinal total e no sinal vocal. Neste caso, um identificador, que é capaz de determinar o tipo de código de codificação com respeito ao sinal total e ao sinal vocal, deve ser concebido em um fluxo de bits, com um decodificador desempenhando o processo de identificação do tipo de código para a determinação do identificador, decodificando os sinais, e então, removendo os componentes vocais. Neste processo, conforme mencionado acima, a soma ou a diferença é utilizada. A informação sobre o identificador pode incluir informação sobre se um sinal residual foi empregado utilizando o mesmo código como aquele para um sinal de decodificação original, o tipo de código utilizado para codificação de um sinal residual, e assim por diante.

Ainda, os diferentes códigos de codificação podem ser empregados para o sinal total e para o sinal vocal. Por exemplo, o sinal vocal (ou seja, o sinal residual) sempre faz uso de um código fixo. Nesta situação, não é necessário um identificador para o sinal residual, e somente um código pré-determinado pode ser empregado para decodificação do sinal total. Entretanto, nesta situação, um processo de remoção do sinal residual do sinal total é restrito a um domínio aonde o processamento entre os dois sinais é possível de maneira imediata, tal como um domínio temporal ou um domínio de sub-faixas. Por exemplo, em um domínio, tal como um mdct, é impossível o processamento imediato entre dois sinais.

Mais ainda, de acordo com a presente invenção, pode ser liberado um sinal de karaokê compreendendo somente de um sinal de objeto de fundo. Pode ser gerado um sinal de canal múltiplo pela execução de um processo adicional de elevação do sinal de karaokê. Por exemplo, caso o MPEG de entorno seja aplicado adicionalmente junto ao sinal de karaokê gerado pela presente invenção, um sinal de karaokê em formato 5,1, pode ser gerado.

Incidentalmente, nas modalidades anteriores, foi-se descrita à quantidade de objetos musicais e objetos principais, ou a quantidade de objetos de fundo e objetos principais dentro de um quadro que se apresentam idênticas. Contudo, a quantidade de objetos musicais e objetos principais, ou a quantidade de objetos de fundo e objetos principais dentro de um quadro podem diferir. Por exemplo, a música pode estar presente em cada quadro e o objeto principal pode estar presente a cada dois quadros. Nesta oportunidade, o objeto principal pode ser decodificado e o resultado da decodificação pode ser aplicado junto a dois quadros.

A música e o objeto principal apresentam diferentes frequências de amostragem. Por exemplo, quando a frequência de amostragem musical é 44,1 kHz e a frequência de amostragem de um objeto principal é 22,05 kHz, os coeficientes MDCT do objeto principal podem ser calculados e a mixagem pode ser então realizada, somente em uma região correspondendo aos coeficientes MDCT musicais. Emprega-se o princípio que o som vocal apresenta uma faixa de frequência mais baixa do que aquela faixa de frequência para o som instrumental com respeito a um sistema karaokê, e isto é algo vantajoso porque pode reduzir a quantidade de dados.

Além do mais, de acordo com a presente invenção, os códigos lidos por um processador podem ser implementados em uma mídia de gravação fixa via processador. A mídia de gravação fixa via processador pode incluir todos os tipos de dispositivos de gravação aonde são armazenados os dados que podem ser lidos pelo processador. Exemplos da mídia de gravação fixa via processador pode incluir fitas magnéticas, ROM, RAM, CD-ROM, discos floppy, dispositivos óticos para armazenagem de dados, e assim por

diante, incluindo também ondas portadoras, tais como, transmissões via Internet. Ainda, a mídia de gravação fixa via processador pode ser distribuída nos sistemas conectados através de uma rede de trabalho, e os códigos lidos pelo processador podem ser armazenados e executados em uma maneira distribuída.

5           Embora a presente invenção tenha sido descrita em conexão com o que se considera presentemente como as modalidades preferidas, se deve compreender que a presente invenção não se faz limitada por estas modalidades específicas, mas são possíveis várias modificações que podem ser concebidas por especialistas na área. Deve-se observar  
10           do emprego da presente invenção.

#### Aplicabilidade Industrial

          A presente invenção pode ser empregada para processos de codificação e decodificação de sinais de áudio orientados a objeto, etc., para o processamento de sinais de objeto com uma associação de base por grupo, e pode fornecer modos de reprodução  
15           tais como, um modo karaokê, um modo solo, e um modo genérico.

## REIVINDICAÇÕES

1. Método de decodificação de áudio, **CARACTERIZADO** pelo fato de compreender:

5 extração de um primeiro sinal de áudio e de um primeiro parâmetro de áudio em que um objeto musical é codificado em uma base de canal e um segundo sinal de áudio e um segundo parâmetro de áudio em que um objeto vocal é codificado em uma base objeto, a partir de um sinal de áudio;

geração de um terceiro sinal de áudio através do emprego de, pelo menos, um dos primeiros e segundos sinais de áudio; e

10 geração de um sinal de áudio de múltiplos canais através do emprego de, pelo menos, um dos primeiros e segundos parâmetros de áudio e o terceiro sinal de áudio.

2. Método de decodificação de áudio, de acordo com a reivindicação 1, **CARACTERIZADO** pelo fato de que o primeiro sinal de áudio é obtido através da codificação de, pelo menos, dois objetos musicais, e o segundo sinal de áudio é obtido pela  
15 codificação de, pelo menos, dois objetos vocais.

3. Método de decodificação de áudio, de acordo com a reivindicação 1, **CARACTERIZADO** pelo fato do terceiro sinal de áudio ser gerado com base em um comando de controle para usuário.

4. Método de decodificação de áudio, de acordo com a reivindicação 1, **CARACTERIZADO** pelo fato do terceiro sinal de áudio ser gerado com base da  
20 adição/subtração de um sinal, pelo menos, de um dos primeiros e segundos sinais de áudio.

5. Método de decodificação de áudio, de acordo com a reivindicação 1, **CARACTERIZADO** pelo fato do terceiro sinal de áudio ser gerado pela remoção de, pelo menos, um dos primeiros e segundos sinais de áudio.

25 6. Método de decodificação de áudio, de acordo com a reivindicação 1, **CARACTERIZADO** pelo fato do primeiro sinal de áudio ser um sinal que não inclui um componente vocal.

7. Método de decodificação de áudio, de acordo com a reivindicação 1, **CARACTERIZADO** pelo fato do sinal de áudio ser um sinal recebido a partir de um sinal de  
30 ampla transmissão.

8. Aparelhagem de decodificação de áudio, **CARACTERIZADA** pelo fato de compreender:

um multiplexador para extrair um sinal de mistura descendente e a informação paralela a partir de um fluxo de bits recebido;

35 um decodificador de objeto para geração de um terceiro sinal de áudio através do emprego de, pelo menos, um primeiro sinal de áudio aonde um objeto musical extraído do sinal de mistura descendente é codificado em uma base de canal, e um segundo sinal de

áudio aonde um objeto vocal extraído do sinal de mistura descendente é codificado em uma base objeto; e

um decodificador de múltiplos canais para geração de um sinal de áudio de múltiplos canais através do emprego de, pelo menos, um primeiro parâmetro de áudio e um segundo parâmetro de áudio extraídos da informação paralela, e terceiro sinal de áudio.

9. Aparelhagem de decodificação de áudio, de acordo com a reivindicação 8, **CARACTERIZADA** pelo fato do decodificador de objeto gerar o terceiro sinal de áudio com base da adição/subtração de um sinal de, pelo menos, um dos primeiros e segundos sinais de áudio.

10. Método de decodificação de áudio, **CARACTERIZADO** pelo fato de compreender as etapas de:

receber um sinal de mistura descendente;

extrair um primeiro sinal de áudio aonde um objeto musical, incluindo um objeto vocal, é codificado e, um segundo sinal de áudio aonde um objeto vocal é codificado, a partir do sinal de mistura descendente; e

gerar qualquer um dentre um sinal de áudio, incluindo somente o objeto vocal, um sinal de áudio compreendendo o objeto vocal, e um sinal de áudio não incluindo o objeto vocal baseado no primeiro e segundo sinais de áudio.

11. Método de decodificação de áudio, de acordo com a reivindicação 10, **CARACTERIZADO** pelo fato do primeiro sinal de áudio ser de um sinal que é codificado em uma base de canal, e o segundo sinal de áudio ser de um sinal que é codificado em uma base objeto.

12. Método de decodificação de áudio, de acordo com a reivindicação 10, **CARACTERIZADO** pelo fato do segundo sinal de áudio ser de um sinal em forma residual.

13. Aparelhagem de decodificação de áudio, **CARACTERIZADA** pelo fato de compreender:

um decodificador de objeto para geração de qualquer um dentre um sinal de áudio incluindo somente um objeto vocal, um sinal de áudio compreendendo o objeto vocal, e um sinal de áudio não incluindo o objeto vocal baseado em um primeiro sinal de áudio aonde um objeto musical extraído de um sinal de mistura descendente é codificado, e um segundo sinal de áudio aonde um objeto vocal extraído do sinal de mistura descendente é codificado; e

um decodificador de múltiplos canais para geração de um sinal de áudio de múltiplos canais através do emprego de uma saída de sinal a partir do decodificador de objeto.

14. Aparelhagem de decodificação de áudio, de acordo com a reivindicação 13, **CARACTERIZADA** pelo fato do primeiro sinal de áudio ser de um sinal que é codificado em

uma base de canal, e o segundo sinal de áudio ser de um sinal que é codificado em uma base objeto.

15. Aparelhagem de decodificação de áudio, de acordo com a reivindicação 13, **CARACTERIZADA** pelo fato de compreender ainda um demultiplexador para extração do sinal de mistura descendente e da informação paralela utilizados para geração do sinal de áudio de múltiplos canais a partir de um fluxo de bits recebido.

16. Método de codificação de áudio, **CARACTERIZADO** pelo fato de compreender as etapas de:

gerar um primeiro sinal de áudio aonde um objeto musical é codificado em uma base de canal, e um primeiro parâmetro de áudio correspondendo ao objeto musical;

gerar um segundo sinal de áudio aonde seja codificado um objeto vocal em uma base objeto, e um segundo parâmetro de áudio correspondendo ao objeto vocal; e

gerar um fluxo de bits incluindo o primeiro e segundo sinais de áudio, e o primeiro e segundo parâmetros de áudio.

17. Aparelhagem de codificação de áudio, **CARACTERIZADA** pelo fato de compreender:

um codificador de múltiplos canais para geração de um primeiro sinal de áudio aonde um objeto musical é codificado em uma base de canal, e um primeiro parâmetro de áudio baseado em canal com respeito ao objeto musical;

um codificador de objeto para geração de um segundo sinal de áudio aonde um objeto vocal é codificado em uma base objeto, e um segundo parâmetro de áudio baseado em objeto com respeito ao objeto vocal; e

um multiplexador para geração de um fluxo de bits incluindo o primeiro e segundo sinais de áudio, e primeiro e segundo parâmetros de áudio.

18. Mídia de gravação, **CARACTERIZADA** pelo fato de um programa para execução de um método de decodificação do tipo definido em qualquer uma das reivindicações de 1 a 7 em um processador ser gravado, a mídia de gravação sendo lida pelo processador.

19. Mídia de gravação, **CARACTERIZADA** pelo fato de um programa para a execução de um método de codificação do tipo definido na reivindicação 16 em um processador ser gravado, a mídia de gravação sendo lida pelo processador.

Fig. 1

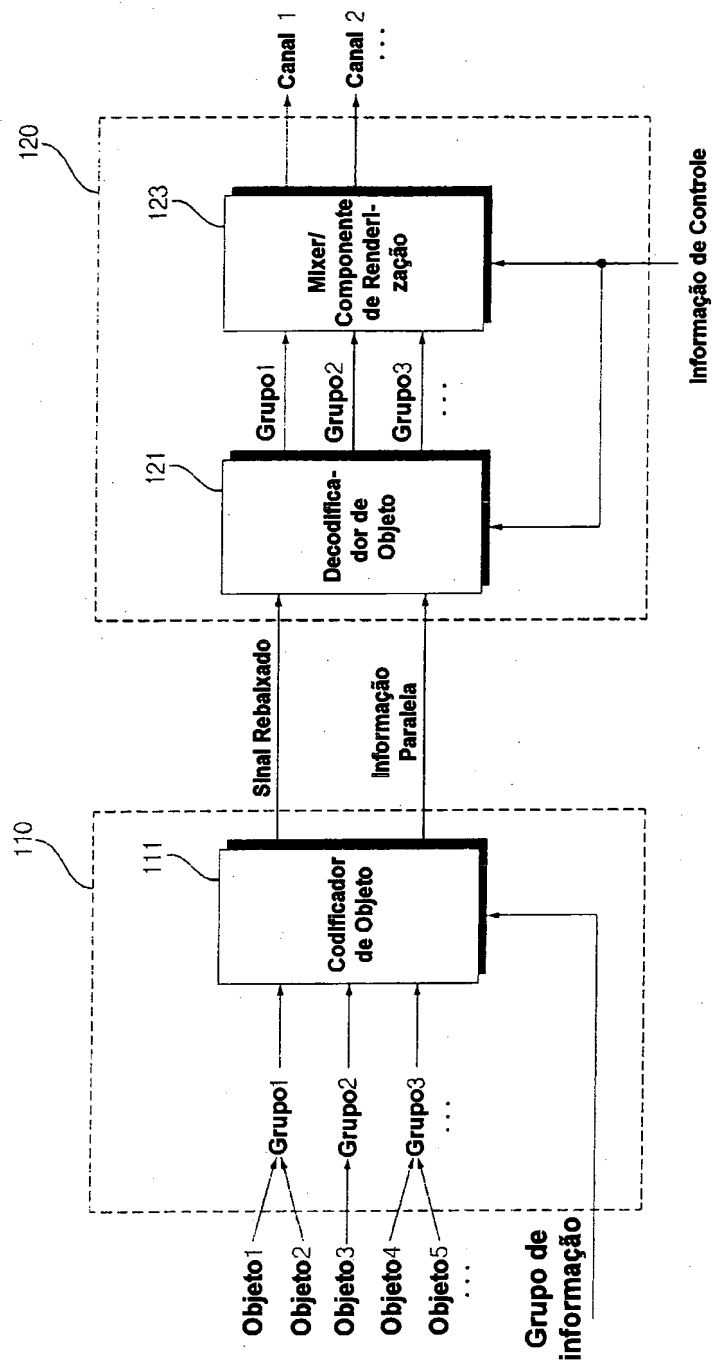


Fig. 2

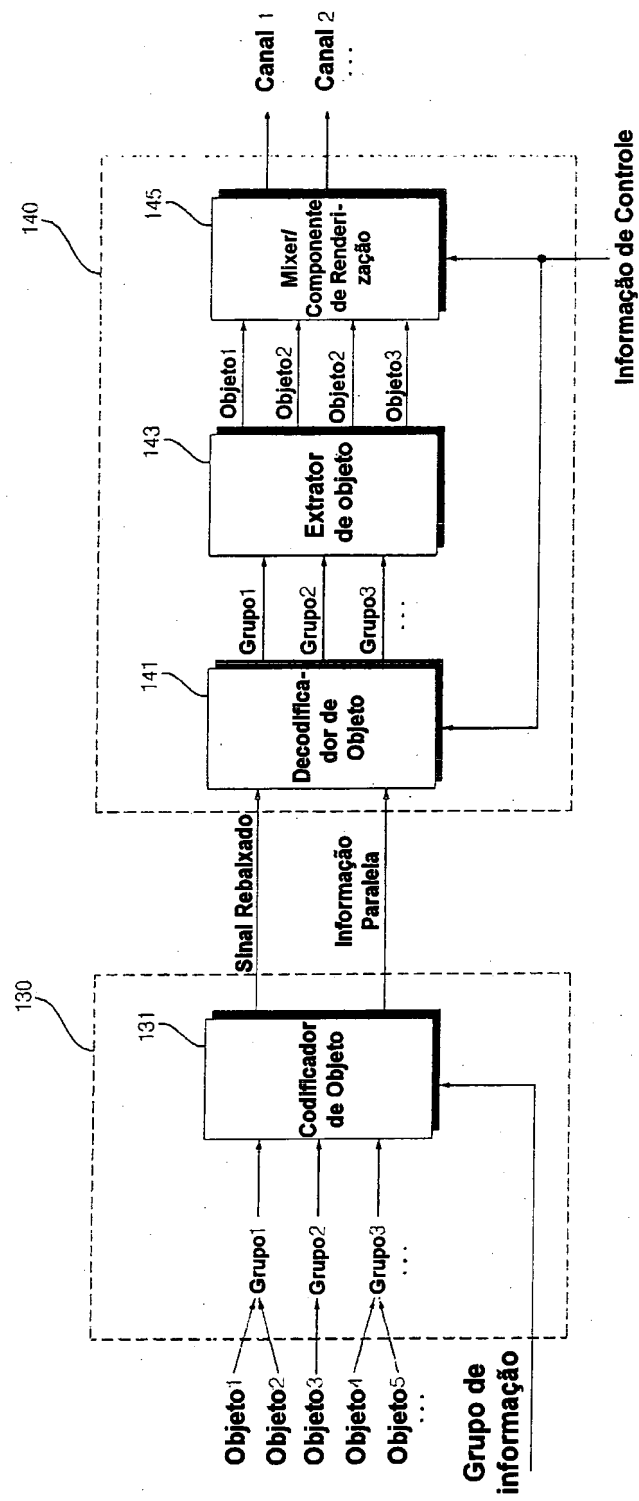


Fig. 3

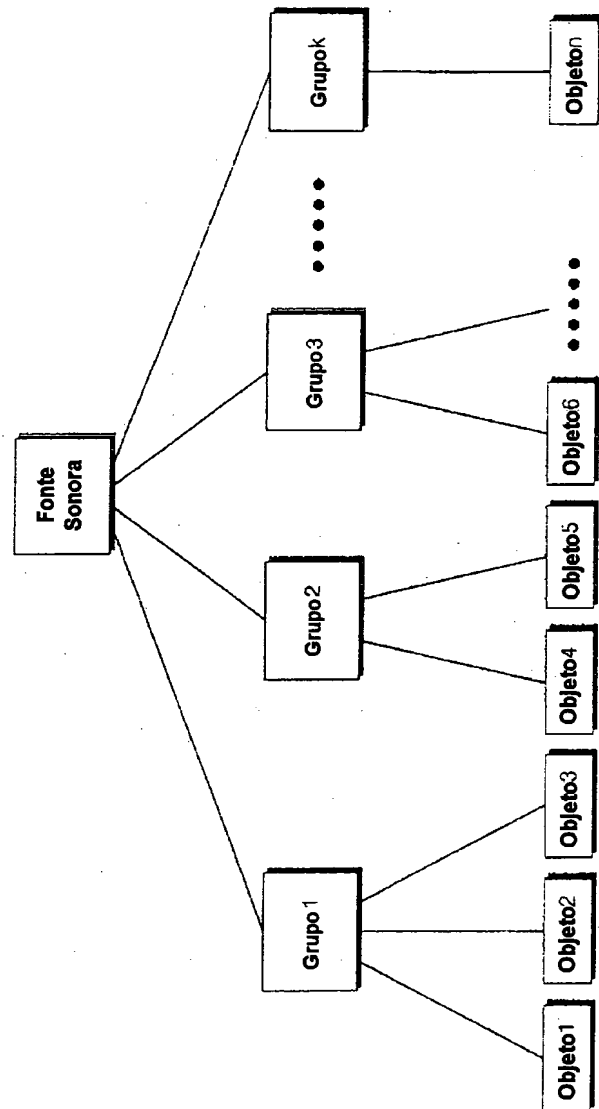


Fig. 4

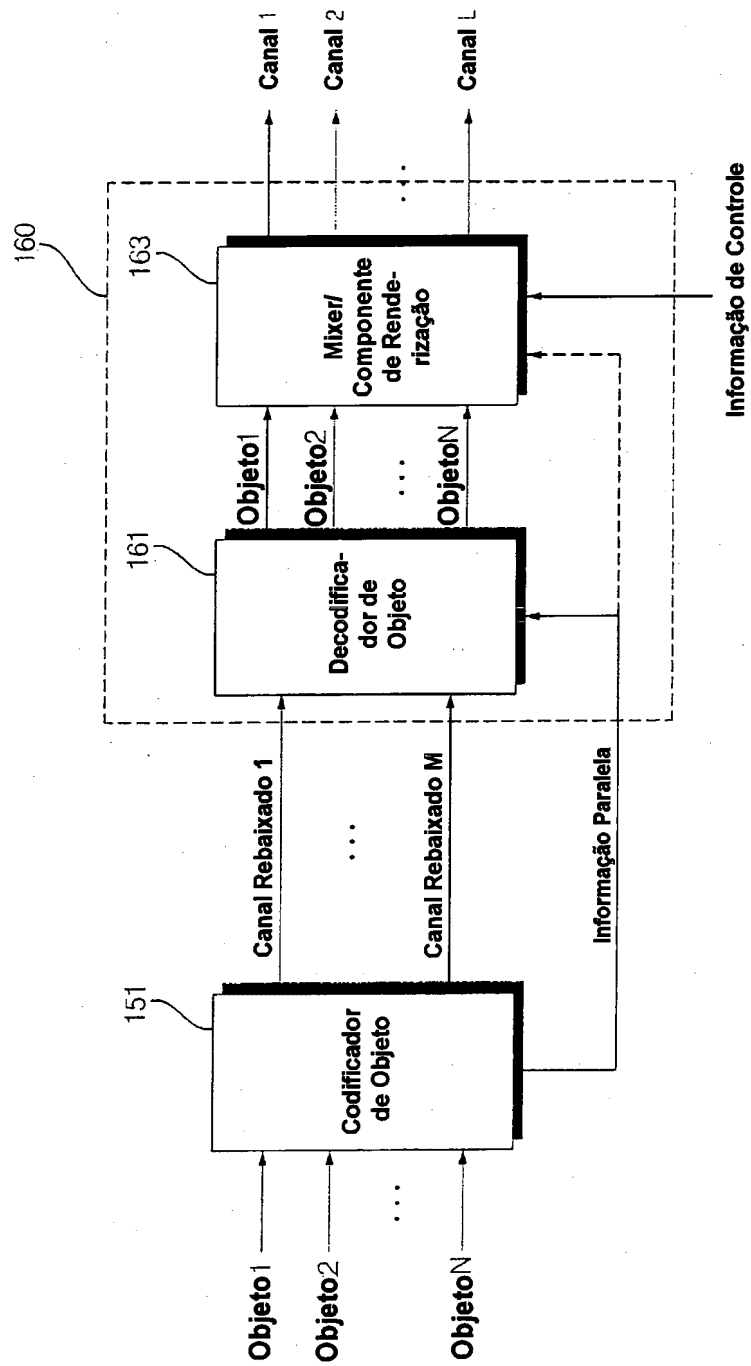
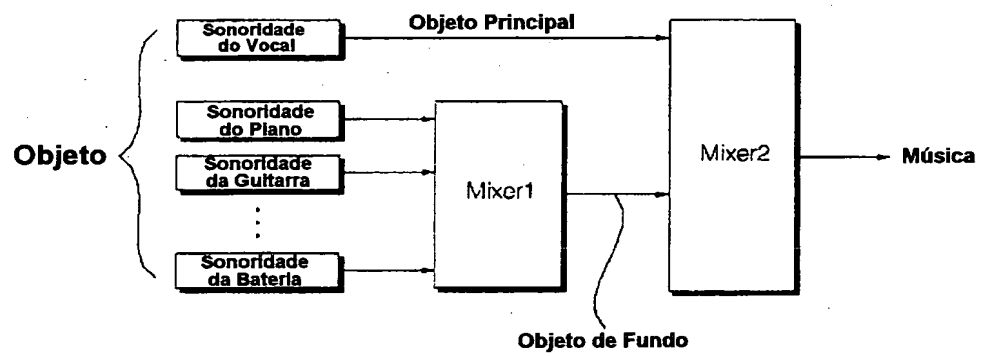
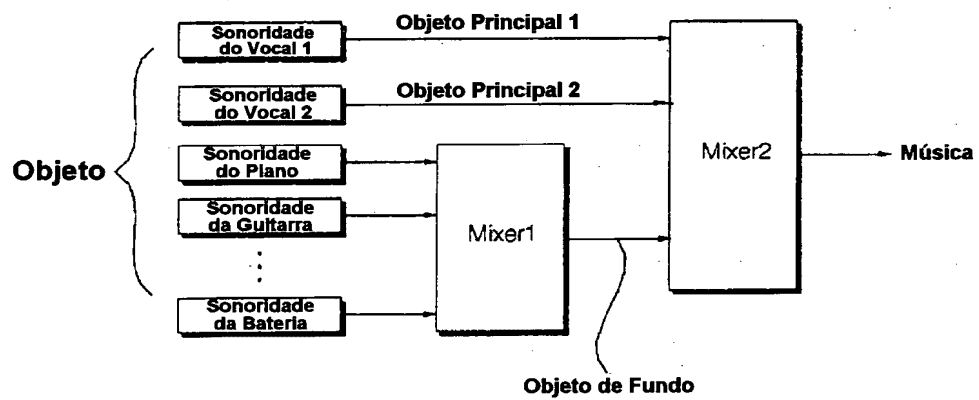


Fig. 5



(a)



(b)

Fig. 6

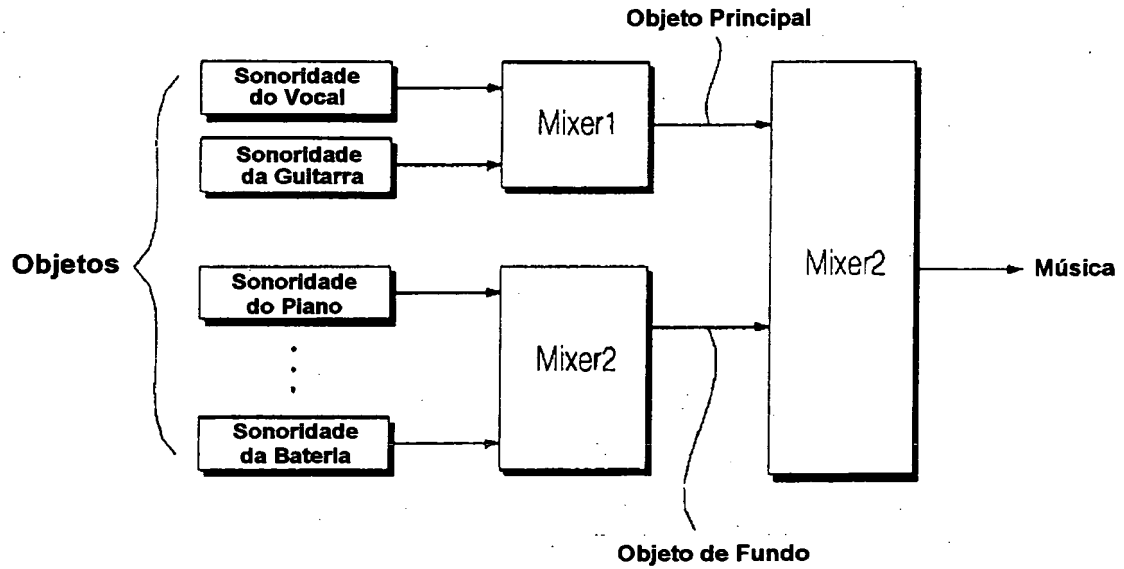


Fig. 7

Formato do fluxo de bits ( unidade de quadro)

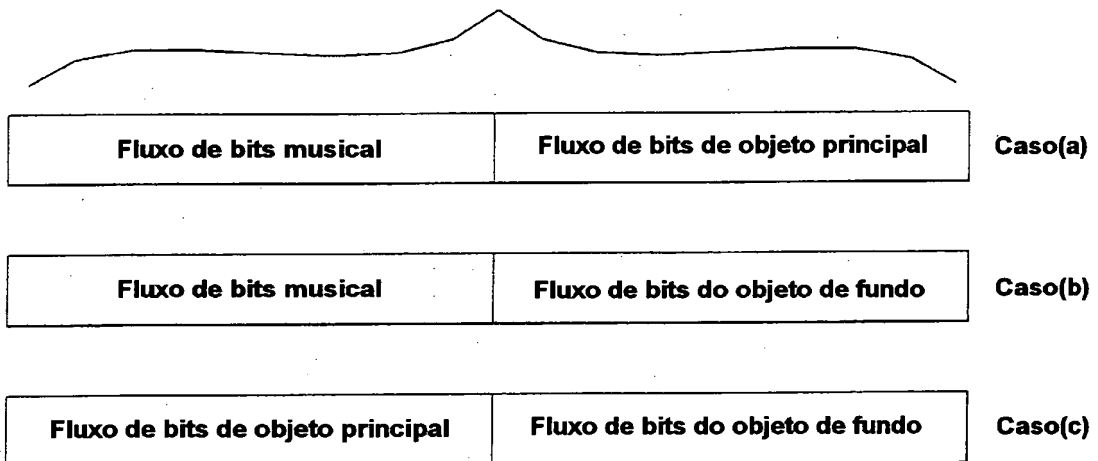


Fig. 8

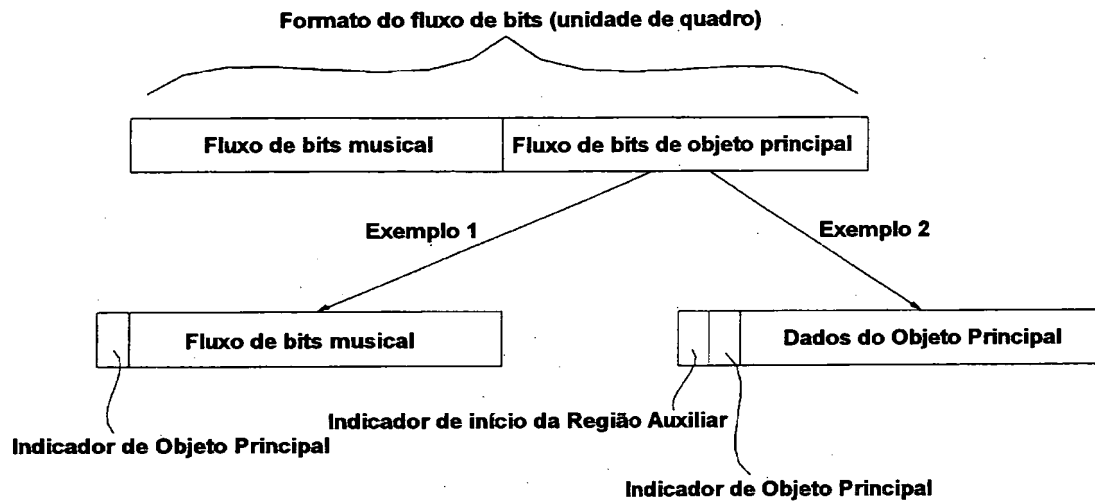


Fig. 9

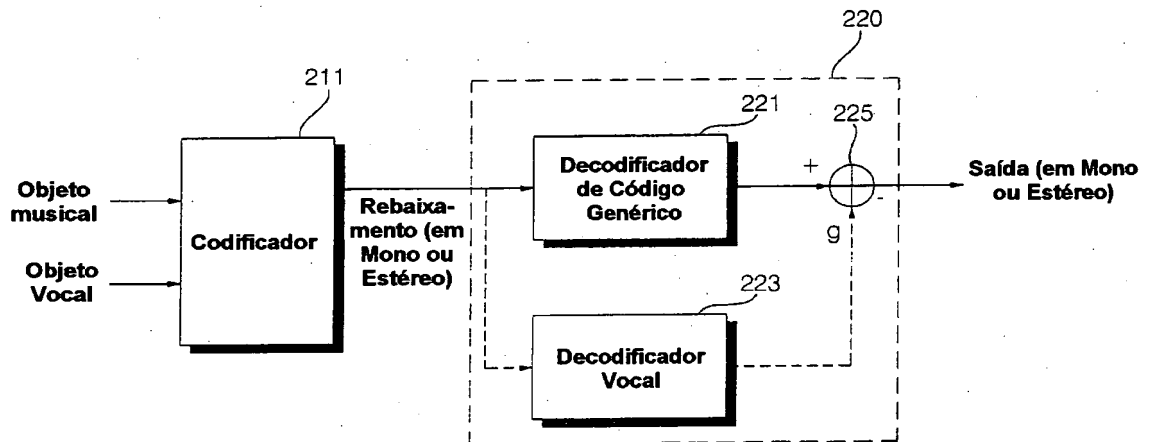


Fig. 10

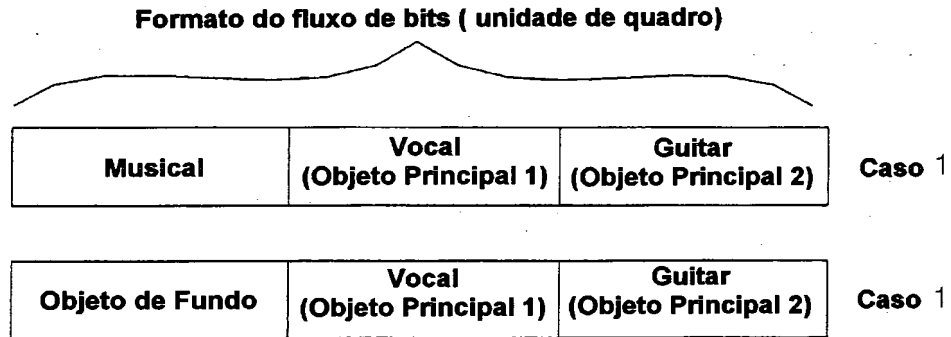


Fig. 11

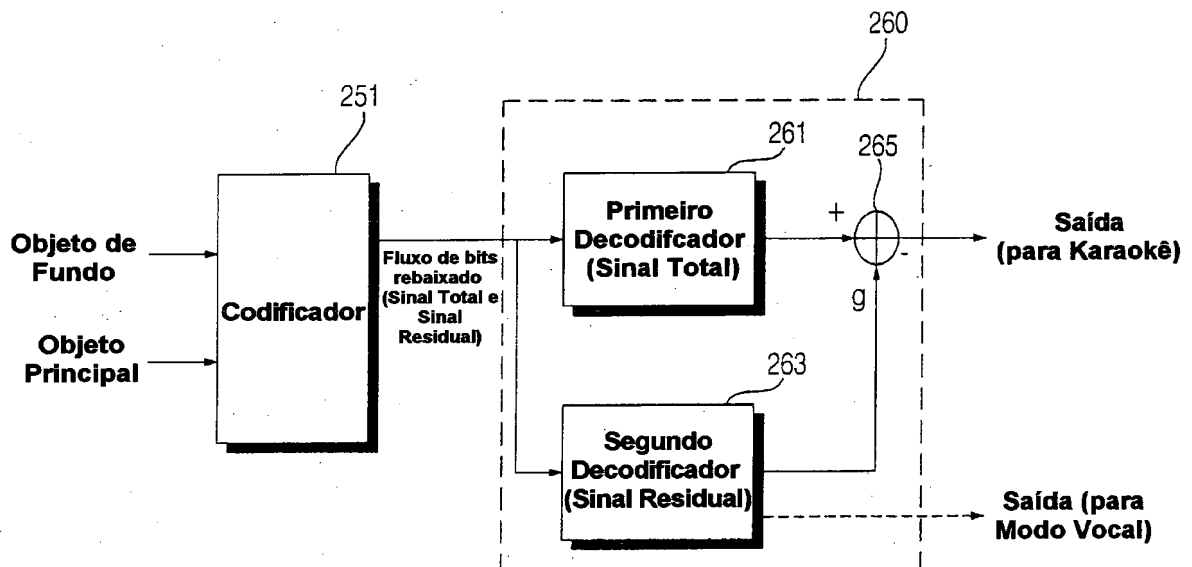


Fig. 12

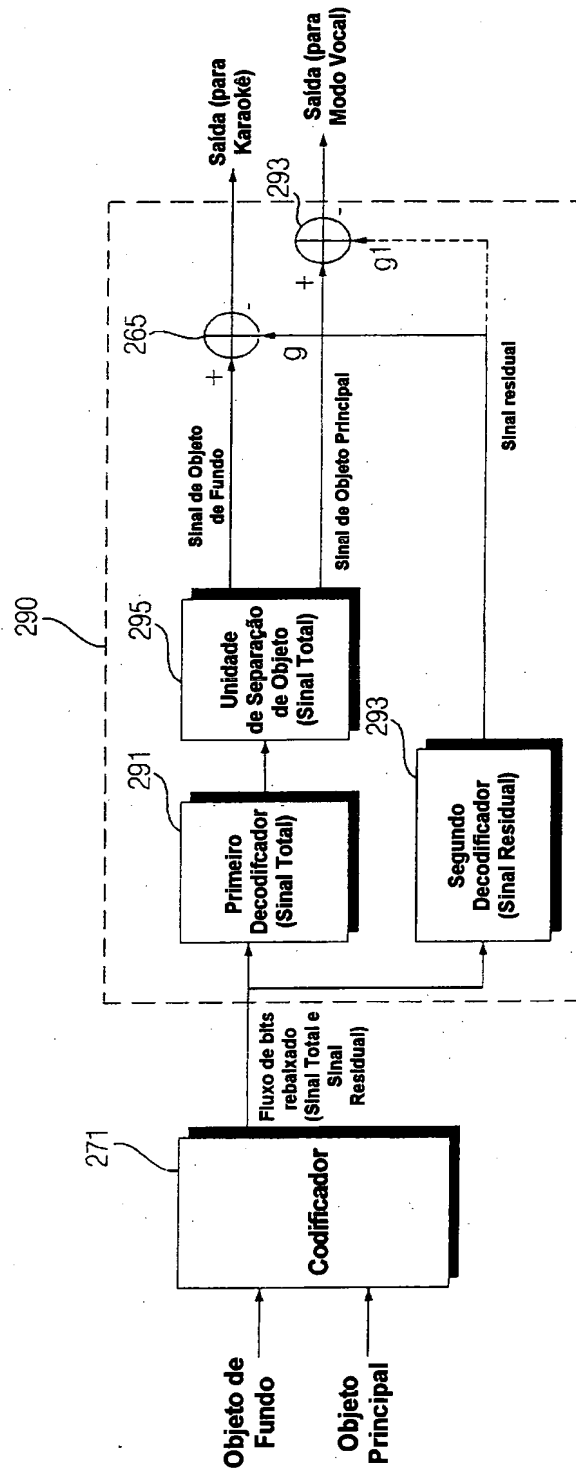


Fig. 13

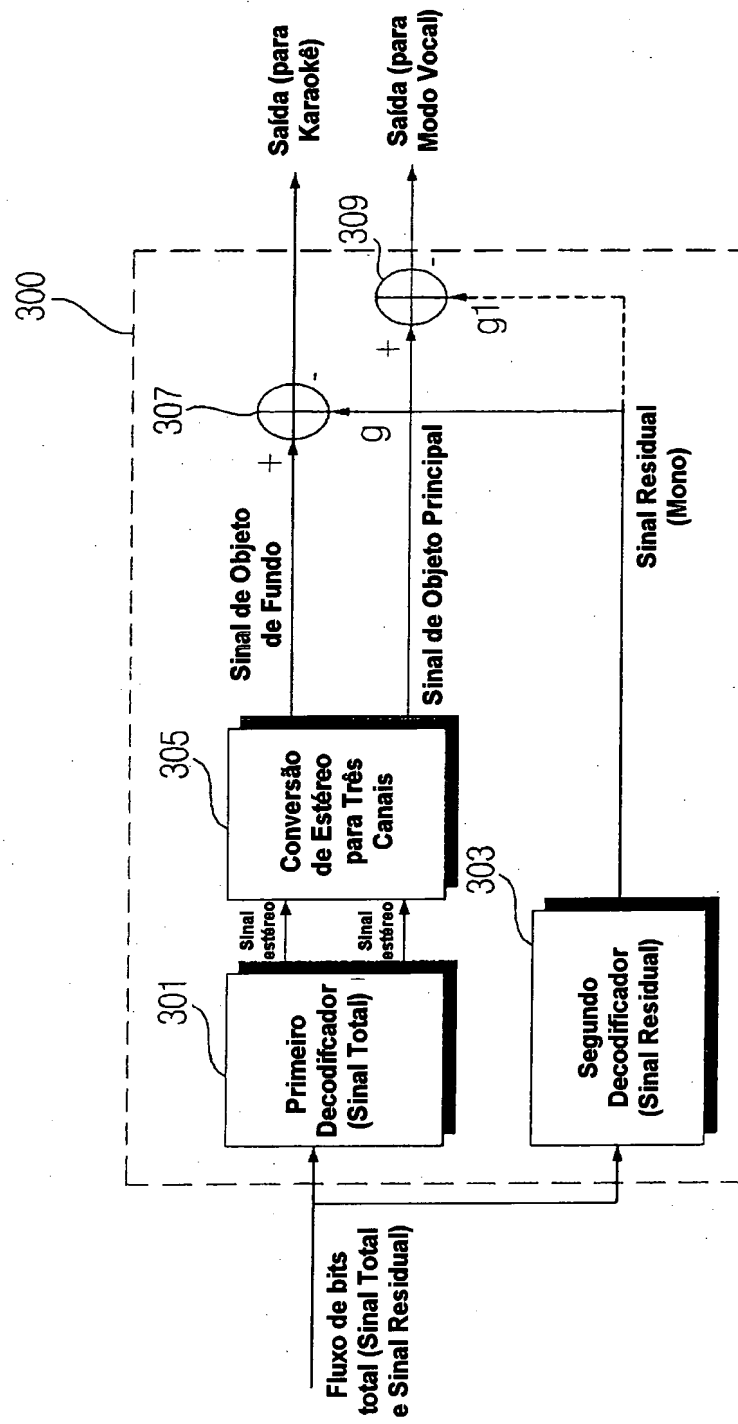


Fig. 14

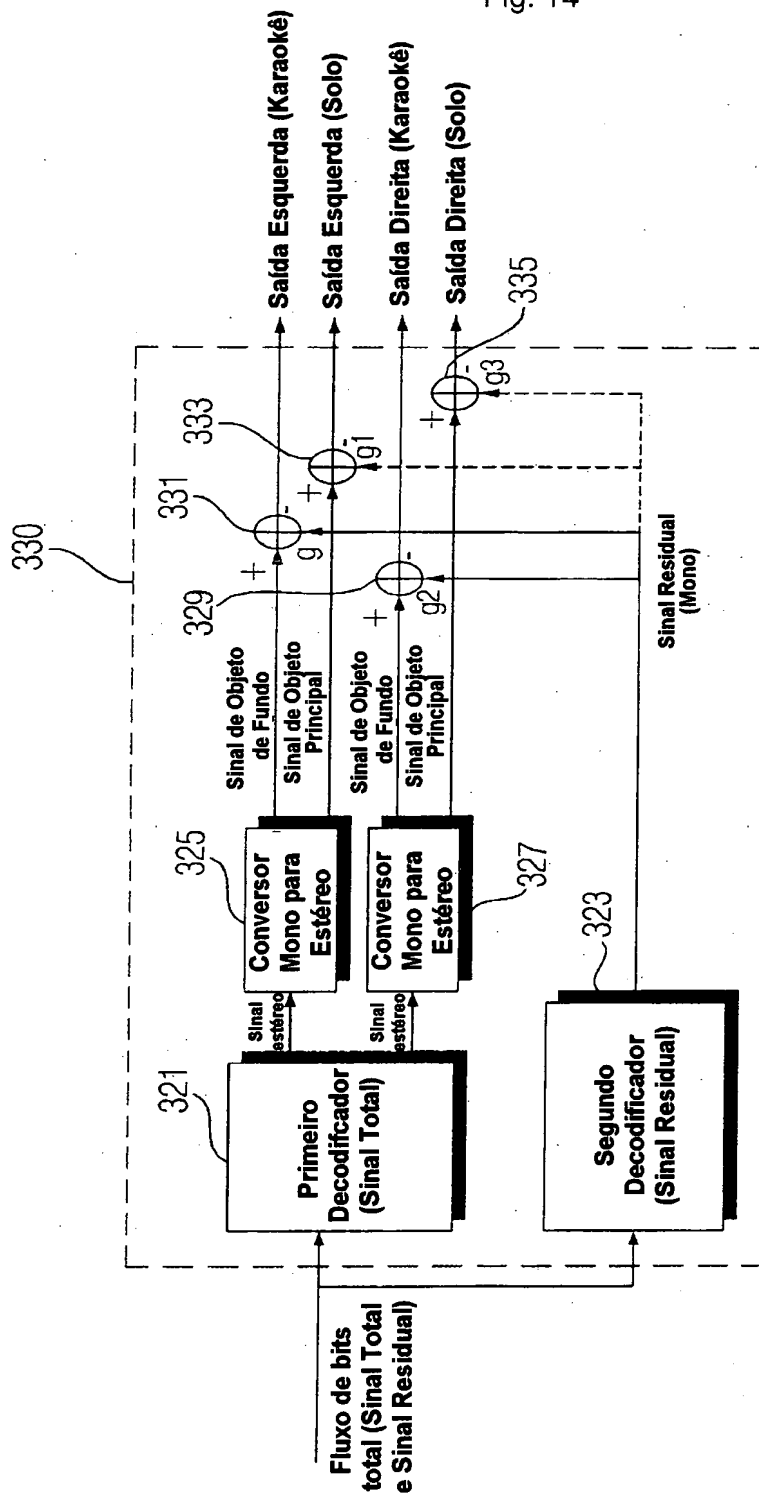


Fig. 15

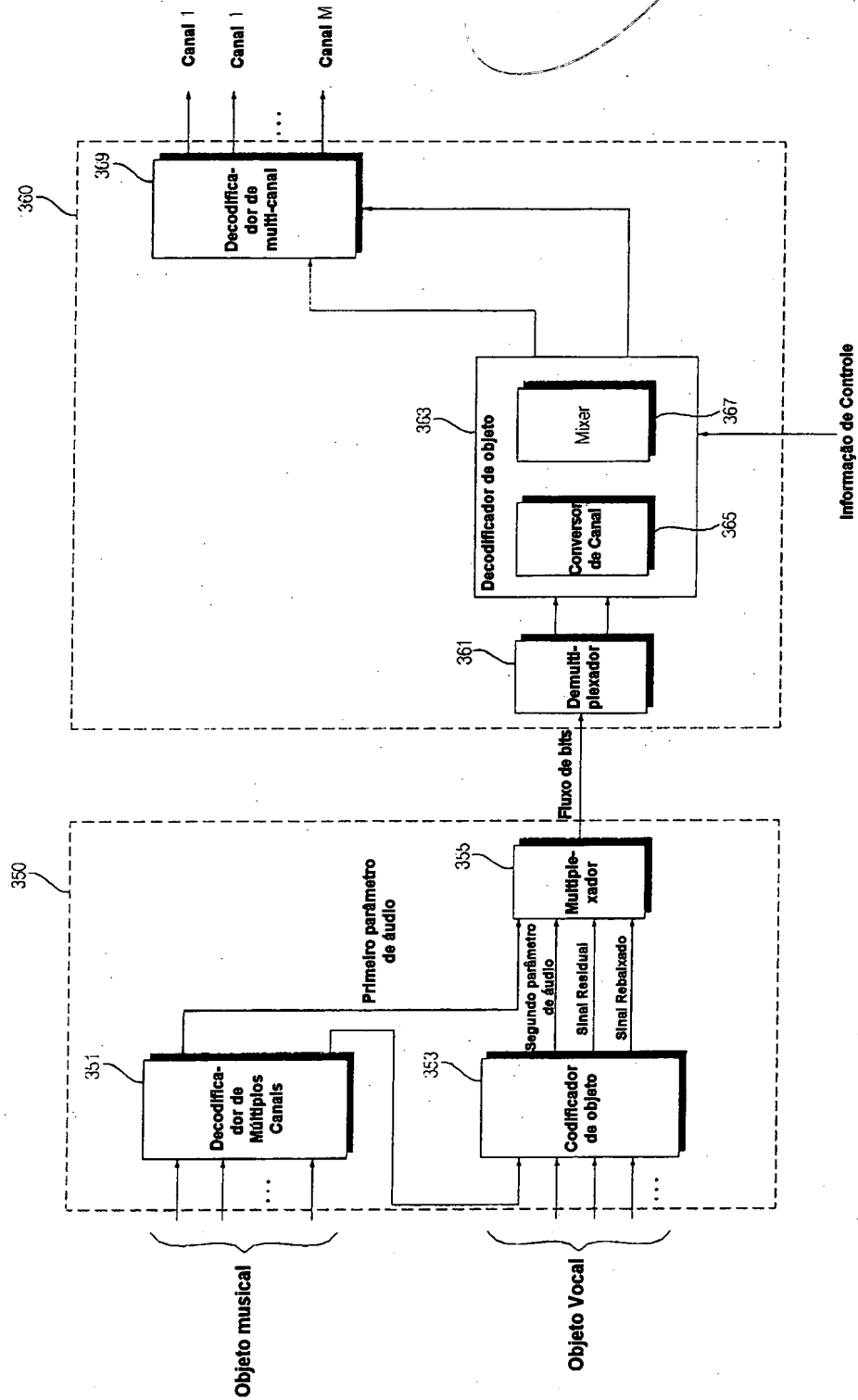
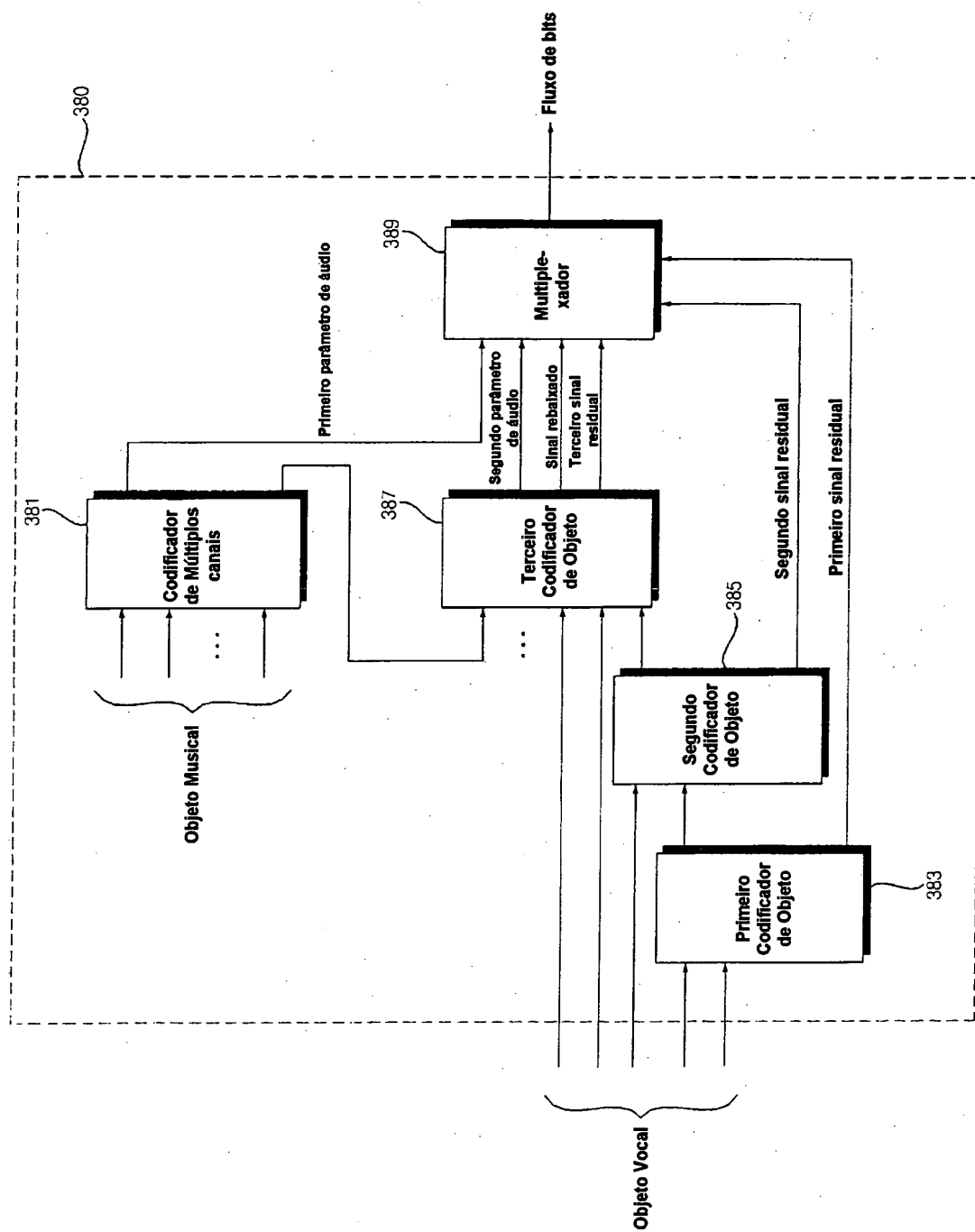


Fig. 16



RESUMO**"MÉTODO PARA CODIFICAÇÃO E DECODIFICAÇÃO DE SINÁL DE ÁUDIO ORIENTADO A OBJETO E APARELHAGEM PARA O MESMO"**

A presente invenção refere-se a um método e aparelhagem para a codificação e decodificação de sinais de áudio orientados a objeto. Este método de decodificação de áudio inclui a extração de um primeiro sinal de áudio e de um primeiro parâmetro de áudio aonde um objeto musical é codificado em uma base de canal, e um segundo sinal de áudio e um segundo parâmetro de áudio aonde um objeto vocal é codificado em uma base objeto, a partir de um sinal de áudio, gerando um terceiro sinal de áudio através do emprego de, pelo menos, um dos primeiros e segundos sinais de áudio, e gerando um sinal de áudio de múltiplos canais através do emprego de, pelo menos, um dos primeiros e segundos parâmetros de áudio e o terceiro sinal de áudio. Resultando que, pode-se reduzir de forma eficiente a quantidade de cálculos nos processos de codificação e decodificação e no tamanho de um fluxo de bits que é codificado.