



ФЕДЕРАЛЬНАЯ СЛУЖБА  
ПО ИНТЕЛЛЕКТУАЛЬНОЙ СОБСТВЕННОСТИ,  
ПАТЕНТАМ И ТОВАРНЫМ ЗНАКАМ

**(12) ОПИСАНИЕ ИЗОБРЕТЕНИЯ К ПАТЕНТУ**

(21), (22) Заявка: 2008140142/09, 14.02.2008

(24) Дата начала отсчета срока действия патента:  
14.02.2008

(30) Конвенционный приоритет:

14.02.2007 US 60/901,089

16.02.2007 US 60/901,642

28.02.2007 US 60/903,818

13.04.2007 US 60/907,689

27.04.2007 US 60/924,027

02.07.2007 US 60/947,620

06.07.2007 US 60/948,373

(43) Дата публикации заявки: 20.04.2010

(45) Опубликовано: 10.12.2010 Бюл. № 34

(56) Список документов, цитированных в отчете о  
поиске: WO 2007004828 A2, 11.01.2007. WO  
2007004830 A1, 11.01.2007. EP 0857375 A1,  
12.08.1998. US 2006002572 A1, 05.01.2006. US  
2003167173 A1, 04.09.2003. WO 2006060279 A1,  
08.06.2006. RU 2004133032 A, 20.04.2005. RU  
2005104123 A, 10.07.2005.(85) Дата перевода заявки РСТ на национальную  
фазу: 09.10.2008(86) Заявка РСТ:  
KR 2008/000885 (14.02.2008)(87) Публикация РСТ:  
WO 2008/100100 (21.08.2008)Адрес для переписки:  
129090, Москва, ул.Б.Спасская, 25, стр.3,  
ООО "Юридическая фирма Городиский и  
Партнеры", пат.пов. Ю.Д.Кузнецову

(72) Автор(ы):

КИМ Донг Соо (KR),  
ПАНГ Хее Сук (KR),  
ЛИМ Дзае Хиун (KR),  
ЙООН Сунг Йонг (KR),  
ЛИ Хиун Коок (KR)

(73) Патентообладатель(и):

ЭлДжи ЭЛЕКТРОНИКС ИНК. (KR)

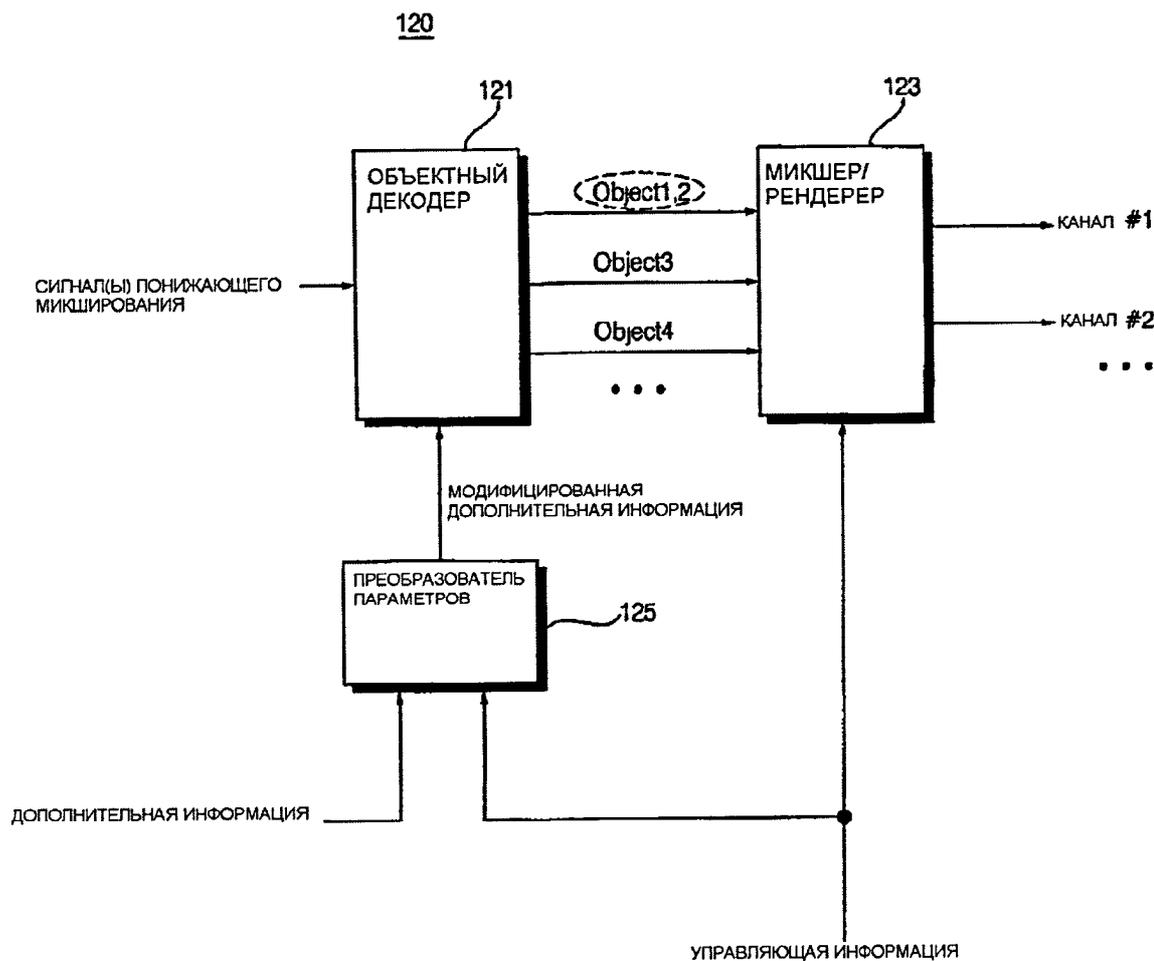
**(54) СПОСОБЫ И УСТРОЙСТВА КОДИРОВАНИЯ И ДЕКОДИРОВАНИЯ  
ОСНОВЫВАЮЩИХСЯ НА ОБЪЕКТАХ ОРИЕНТИРОВАННЫХ АУДИОСИГНАЛОВ**

(57) Реферат:

Изобретение относится к способам и  
устройствам кодирования и декодирования  
многоканального аудио, в которых операциикодирования и декодирования производятся  
над объектно-базированными аудиосигналами,  
Техническим результатом является обеспечение  
эффективности обработки основывающихся на

объектах аудиосигналов. Указанный технический результат достигается тем, что способ декодирования аудио включает в себя прием сигнала понижающего микширования, основывающейся на объектах дополнительной информации и управляющей информации, при этом сигнал понижающего микширования содержит, по меньшей мере, два канальных сигнала понижающего микширования и сгенерирован посредством понижающего микширования множества объектных сигналов, а управляющая информация управляет позицией или уровнем объектного сигнала, включенного в сигнал понижающего микширования; извлечение из основывающейся

на объектах дополнительной информации информации соотношения распределения объектного сигнала между каналами; формирование информации предварительной обработки для модификации канальных сигналов понижающего микширования на основе информации соотношения распределения между каналами и управляющей информации; формирование многоканального аудиосигнала на основе основывающейся на объектах дополнительной информации и управляющей информации и модифицированных канальных сигналов понижающего микширования. 6 н. и 5 з.п. ф-лы, 33 ил.



RU 2406166 C2

RU 2406166 C2



FEDERAL SERVICE  
FOR INTELLECTUAL PROPERTY,  
PATENTS AND TRADEMARKS

**(12) ABSTRACT OF INVENTION**

(21), (22) Application: **2008140142/09, 14.02.2008**

(24) Effective date for property rights:  
**14.02.2008**

(30) Priority:  
14.02.2007 US 60/901,089  
16.02.2007 US 60/901,642  
28.02.2007 US 60/903,818  
13.04.2007 US 60/907,689  
27.04.2007 US 60/924,027  
02.07.2007 US 60/947,620  
06.07.2007 US 60/948,373

(43) Application published: **20.04.2010**

(45) Date of publication: **10.12.2010 Bull. 34**

(85) Commencement of national phase: **09.10.2008**

(86) PCT application:  
**KR 2008/000885 (14.02.2008)**

(87) PCT publication:  
**WO 2008/100100 (21.08.2008)**

Mail address:  
**129090, Moskva, ul.B.Spaskaja, 25, str.3, OOO  
"Juridicheskaja firma Gorodisskij i Partnery",  
pat.pov. Ju.D.Kuznetsovu**

(72) Inventor(s):

**KIM Dong Soo (KR),  
PANG Khee Suk (KR),  
LIM Dzae Khiun (KR),  
JOON Sung Jong (KR),  
LI Khiun Kook (KR)**

(73) Proprietor(s):

**EhIDzhi EhLEKTRONIKS INK. (KR)**

**(54) CODING AND DECODING METHODS AND DEVICES BASED ON OBJECTS OF ORIENTED AUDIO SIGNALS**

(57) Abstract:

FIELD: information technology.

SUBSTANCE: audio decoding method involves receiving a step-down mixing signal based on additional information objects and control information, where the step-down mixing signal contains at least two step-down mixing channel signals and is generated through step-down mixing of multiple object signals, and control information controls the position or level or the object signal contained in the step-down mixing signal; extraction of the ratio of distribution of the object signals

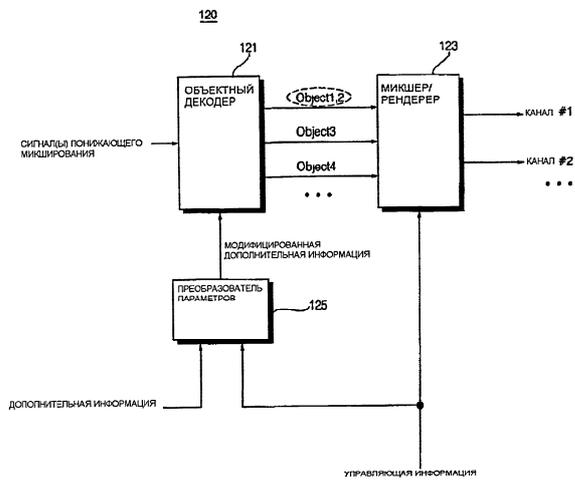
between channels from the object-based additional information; generation of pre-processing information for modifying step-down mixing channel signals based on information on the ratio of distribution between channels and control information; generation of a mutichannel audio signal based on the object-based additional information and control information and modified step-down mixing channel signals.

EFFECT: efficient processing of object-based audio signals.

11 cl, 33 dwg

RU 2 406 166 C2

RU 2 406 166 C2



ФИГ.2

RU 2406166 C2

RU 2406166 C2

**Область техники, к которой относится изобретение**

Настоящее изобретение относится к способу и устройству кодирования аудио и способу и устройству декодирования аудио, в которых основывающиеся на объектах аудиосигналы могут быть эффективно обработаны посредством выполнения операций кодирования и декодирования.

**Предшествующий уровень техники**

В общем, в методиках кодирования и декодирования многоканального аудио в отношении некоторого количества канальных сигналов многоканального сигнала выполняется понижающее микширование с получением меньшего количества канальных сигналов, дополнительная информация, касающаяся исходных канальных сигналов, передается, и многоканальный сигнал, имеющий столько же каналов, что и исходный многоканальный сигнал, восстанавливается.

Методики кодирования и декодирования основывающегося на объектах аудио, по сути, аналогичны методикам кодирования и декодирования многоканального аудио в отношении понижающего микширования нескольких источников звука в меньшее число сигналов источника звука и передачи дополнительной информации, касающейся исходных источников звука. Тем не менее, в методиках кодирования и декодирования основывающегося на объектах аудио объектные сигналы, которые являются базовыми сигналами (к примеру, музыкальный инструмент или человеческий голос) канального сигнала, интерпретируются так же, как канальные сигналы в методиках кодирования и декодирования многоканального аудио, и тем самым могут быть кодированы.

Другими словами, в методиках кодирования и декодирования основывающегося на объектах аудио объектные сигналы считаются объектами, которые должны быть кодированы. В этом смысле методика кодирования и декодирования основывающегося на объектах аудио отличаются от методик кодирования и декодирования многоканального аудио, в которых операция кодирования многоканального аудио выполняется просто на основе межканальной информации независимо от числа элементов канального сигнала, которые должны быть кодированы.

**Сущность изобретения*****Техническая задача***

Настоящее изобретение предоставляет способ и устройство кодирования аудио и способ и устройство декодирования аудио, в которых аудиосигналы могут быть кодированы или декодированы так, чтобы аудиосигналы могли быть применены к различным вариантам осуществления.

***Техническое решение***

Согласно аспекту настоящего изобретения предусмотрен способ декодирования аудио, включающий в себя: прием сигнала понижающего микширования и основывающейся на объектах дополнительной информации, при этом сигнал понижающего микширования включает в себя, по меньшей мере, два канальных сигнала понижающего микширования; извлечение информации усиления из основывающейся на объектах дополнительной информации и формирование информации модификации для модификации канальных сигналов понижающего микширования на поканальной основе на базе информации усиления; и модификацию канальных сигналов понижающего микширования посредством применения информации модификации к канальным сигналам понижающего микширования.

Согласно другому аспекту настоящего изобретения предусмотрен способ

кодирования аудио, включающий в себя: формирование сигнала понижающего микширования посредством понижающего микширования объектного сигнала, при этом сигнал понижающего микширования включает в себя, по меньшей мере, два канальных сигнала понижающего микширования; извлечение связанной с объектами информации, касающейся объектного сигнала, и формирование основывающейся на объектах дополнительной информации на основе связанной с объектами информации; и вставку информации усиления для модификации канальных сигналов понижающего микширования на поканальной основе в основывающуюся на объектах дополнительную информацию.

Согласно другому аспекту настоящего изобретения предусмотрено устройство декодирования аудио, включающее в себя: демультимплексор, выполненный с возможностью извлекать сигнал понижающего микширования и основывающуюся на объектах дополнительную информацию из входного аудиосигнала, при этом сигнал понижающего микширования включает в себя, по меньшей мере, два канальных сигнала понижающего микширования; и транскодер, выполненный с возможностью формировать информацию модификации для модификации канальных сигналов понижающего микширования на поканальной основе на базе информации усиления, извлеченной из основывающейся на объектах дополнительной информации, и модифицировать канальные сигналы понижающего микширования посредством применения информации модификации к канальным сигналам понижающего микширования.

Согласно другому аспекту настоящего изобретения предусмотрен машиночитаемый носитель записи, на котором записана компьютерная программа для выполнения способа декодирования аудио, при этом способ декодирования аудио включает в себя: прием сигнала понижающего микширования и основывающейся на объектах дополнительной информации, при этом сигнал понижающего микширования включает в себя, по меньшей мере, два канальных сигнала понижающего микширования; извлечение информации усиления из основывающейся на объектах дополнительной информации и формирование информации модификации для модификации канальных сигналов понижающего микширования на поканальной основе на базе информации усиления; и модификацию канальных сигналов понижающего микширования посредством применения информации модификации к канальным сигналам понижающего микширования.

Согласно другому аспекту настоящего изобретения предусмотрен машиночитаемый носитель записи, на котором записана компьютерная программа для выполнения способа кодирования аудио, при этом способ кодирования аудио включает в себя: формирование сигнала понижающего микширования посредством понижающего микширования объектного сигнала, при этом сигнал понижающего микширования включает в себя, по меньшей мере, два канальных сигнала понижающего микширования; извлечение связанной с объектами информации, касающейся объектного сигнала, и формирование основывающейся на объектах дополнительной информации на основе связанной с объектами информации; и вставку информации усиления для модификации канальных сигналов понижающего микширования на поканальной основе в основывающуюся на объектах дополнительную информацию.

#### **Краткое описание чертежей**

Фиг.1 иллюстрирует блок-схему типичной системы кодирования/декодирования основывающегося на объектах аудио;

Фиг.2 иллюстрирует блок-схему устройства декодирования аудио согласно первому варианту осуществления настоящего изобретения;

Фиг.3 иллюстрирует блок-схему устройства декодирования аудио согласно второму варианту осуществления настоящего изобретения;

5 Фиг.4 иллюстрирует блок-схему устройства декодирования аудио согласно третьему варианту осуществления настоящего изобретения;

Фиг.5 иллюстрирует блок-схему модуля произвольного усиления понижающего микширования (ADG), который может быть использован в модуле декодирования аудио, проиллюстрированном на фиг.4;

10 Фиг.6 иллюстрирует блок-схему устройства декодирования аудио согласно четвертому варианту осуществления настоящего изобретения;

Фиг.7 иллюстрирует блок-схему устройства декодирования аудио согласно пятому варианту осуществления настоящего изобретения;

15 Фиг.8 иллюстрирует блок-схему устройства декодирования аудио согласно шестому варианту осуществления настоящего изобретения;

Фиг.9 иллюстрирует блок-схему устройства декодирования аудио согласно седьмому варианту осуществления настоящего изобретения;

20 Фиг.10 иллюстрирует блок-схему устройства декодирования аудио согласно восьмому варианту осуществления настоящего изобретения;

Фиг.11 и 12 иллюстрируют схемы для пояснения работы транскодера;

Фиг.13-16 иллюстрируют схемы для пояснения конфигурирования основывающейся на объектах дополнительной информации;

25 Фиг.17-22 иллюстрируют схемы для пояснения объединения множества фрагментов основывающейся на объектах дополнительной информации в один фрагмент дополнительной информации;

Фиг.23-27 иллюстрируют схемы для пояснения операции предварительной обработки; и

30 Фиг.28-33 - схемы, иллюстрирующие случай комбинирования множества потоков битов, декодированных с помощью основывающихся на объектах сигналов, в один поток битов.

### **Оптимальный режим осуществления изобретения**

35 Далее настоящее изобретение описывается более подробно со ссылкой на прилагаемые чертежи, на которых показаны примерные варианты осуществления изобретения.

Способ и устройство кодирования аудио и способ и устройство декодирования аудио согласно настоящему изобретению могут быть применены к операциям обработки основывающегося на объектах аудио, но настоящее изобретение не ограничено этим. Другими словами, способ и устройство кодирования аудио и способ и устройство декодирования аудио могут быть применены к различным операциям обработки сигналов, отличным от операций обработки основывающегося на объектах аудио.

45 Фиг.1 иллюстрирует блок-схему типичной системы кодирования/декодирования основывающегося на объектах аудио. В общем, аудиосигналы, вводимые в устройство кодирования основывающегося на объектах аудио, не соответствуют каналам многоканального сигнала, а являются независимыми объектными сигналами. В этом смысле, устройство кодирования основывающегося на объектах аудио отличается от устройства кодирования многоканального аудио, в которое вводятся каналные сигналы многоканального сигнала.

Например, каналные сигналы, такие как сигнал переднего левого канала и сигнал переднего правого канала для 5.1-канального сигнала, могут быть введены в многоканальный аудиосигнал, тогда как объектные сигналы, такие как человеческий голос или звук музыкального инструмента (к примеру, звук скрипки или пианино),  
5 которые являются меньшими объектами, чем каналные сигналы, могут быть введены в устройство кодирования основывающегося на объектах аудио.

Ссылаясь на фиг.1, система кодирования/декодирования основывающегося на объектах аудио включает в себя устройство кодирования основывающегося на объектах аудио и устройство декодирования основывающегося на объектах аудио.  
10 Устройство кодирования основывающегося на объектах аудио включает в себя объектный кодер 100, а устройство декодирования основывающегося на объектах аудио включает в себя объектный декодер 111 и микшер/рендерер 113.

Объектный кодер 100 принимает N объектных сигналов и формирует основывающийся на объектах сигнал понижающего микширования с одним или более каналами и дополнительной информацией, включающей в себя ряд фрагментов информации, извлеченных из N объектных сигналов, таких как информация разности энергии, информация разности фаз и корреляционная информация. Дополнительная информация и основывающийся на объектах сигнал понижающего микширования  
20 объединяются в один поток битов, и поток битов передается в основывающееся на объектах устройство декодирования.

Дополнительная информация может включать в себя флаг, указывающий то, следует выполнять кодирование основывающегося на каналах аудио или кодирование основывающегося на объектах аудио, и тем самым может быть определено, следует выполнять кодирование основывающегося на каналах аудио или кодирование основывающегося на объектах аудио, на основе флага дополнительной информации. Дополнительная информация также может включать в себя информацию энергии, информацию группировки, информацию периода молчания, информацию усиления  
30 понижающего микширования и информацию задержки, касающуюся объектных сигналов.

Дополнительная информация и основывающийся на объектах сигнал понижающего микширования могут быть объединены в один поток битов, и поток битов может передаваться в устройство декодирования основывающегося на объектах аудио.  
35

Объектный декодер 111 принимает основывающийся на объектах сигнал понижающего микширования и дополнительную информацию из устройства кодирования, основывающегося на объектах аудио, и восстанавливает объектные сигналы, имеющие свойства, аналогичные свойствам N объектных сигналов, на основе основывающегося на объектах сигнала понижающего микширования и дополнительной информации. Объектные сигналы, формируемые посредством объектного декодера 111, еще не назначены какой-либо позиции в многоканальном пространстве. Таким образом, микшер/рендерер 113 назначает каждый из объектных  
40 сигналов, сформированных посредством объектного декодера 111, предварительно определенной позиции в многоканальном пространстве и определяет уровни объектных сигналов так, что объектные сигналы могут быть воспроизведены из надлежащих соответствующих позиций, указанных посредством микшера/рендерера 113, с надлежащими соответствующими уровнями,  
45 определенными посредством микшера/рендерера 113. Управляющая информация, касающаяся каждого из объектных сигналов, сформированных посредством объектного декодера 111, может варьироваться во времени, и тем самым  
50

пространственные позиции и уровни объектных сигналов, сформированных посредством объектного декодера 111, могут варьироваться согласно управляющей информации.

5 Фиг.2 иллюстрирует блок-схему устройства 120 декодирования аудио согласно первому варианту осуществления настоящего изобретения. Ссылаясь на фиг.2, устройство 120 декодирования аудио может иметь возможность выполнять адаптивное декодирование посредством анализа управляющей информации.

10 Ссылаясь на фиг.2, устройство 120 декодирования аудио включает в себя объектный декодер 121, микшер/рендерер 123 и преобразователь 125 параметров. Устройство 120 декодирования аудио также может включать в себя демультимплексор (не показан), который извлекает сигнал понижающего микширования и дополнительную информацию из потока битов, вводимого в него, и он применяется  
15 ко всем устройствам декодирования аудио согласно другим вариантам осуществления настоящего изобретения.

Объектный декодер 121 формирует ряд объектных сигналов на основе сигнала понижающего микширования и модифицированной дополнительной информации, предоставленной посредством преобразователя 125 параметров. Микшер/рендерер 123  
20 назначает каждый из объектных сигналов, сформированных посредством объектного декодера 121, предварительно определенной позиции в многоканальном пространстве и определяет уровни объектных сигналов, сформированных посредством объектного декодера 121, согласно управляющей информации. Преобразователь 125 параметров формирует модифицированную дополнительную информацию посредством  
25 комбинирования дополнительной информации и управляющей информации. Затем преобразователь 125 параметров передает модифицированную дополнительную информацию в объектный декодер 121.

Объектный декодер 121 может иметь возможность выполнять адаптивное  
30 декодирование посредством анализа управляющей информации в модифицированной дополнительной информации.

Например, если управляющая информация указывает то, что первый объектный сигнал и второй объектный сигнал назначены одной позиции в многоканальном пространстве и имеют одинаковый уровень, типичное устройство декодирования  
35 аудио может декодировать первый и второй объектные сигналы отдельно, а затем компоновать их в многоканальном пространстве посредством операции микширования/рендеринга.

С другой стороны, объектный декодер 121 устройства 120 декодирования аудио  
40 узнает из управляющей информации в модифицированной дополнительной информации то, что первый и второй объектные сигналы назначены одной позиции в многоканальном пространстве и имеют одинаковый уровень, как если бы они были одним источником звука. Соответственно, объектный декодер 121 декодирует первый и второй объектные сигналы посредством интерпретации их как одного источника  
45 звука без отдельного их декодирования. Как результат, сложность декодирования снижается. Помимо этого, вследствие уменьшения числа источников звука, которые должны быть обработаны, сложность микширования/рендеринга также снижается.

Устройство 120 декодирования аудио может быть эффективно использовано, когда  
50 число объектных сигналов больше числа выходных каналов, поскольку множество объектных сигналов с большой вероятностью должны назначаться одной пространственной позиции.

Альтернативно, устройство 120 декодирования аудио может быть использовано,

когда первый объектный сигнал и второй объектный сигнал назначаются одной позиции в многоканальном пространстве, но имеют различные уровни. В этом случае, устройство 120 декодирования аудио декодирует первый и второй объектные сигналы посредством интерпретации первого и второго объектных сигналов как одного сигнала, вместо декодирования первого и второго объектных сигналов отдельно и передачи декодированных первого и второго объектных сигналов в микшер/рендерер 123. Более конкретно, объектный декодер 121 может получать информацию, касающуюся разности между уровнями первого и второго объектных сигналов, из управляющей информации в модифицированной дополнительной информации, и декодировать первый и второй объектные сигналы на основе полученной информации. Как результат, даже если первый и второй объектные сигналы имеют различные уровни, первый и второй объектные сигналы могут быть декодированы, как если бы они являлись одним источником звука.

Еще альтернативно, объектный декодер 121 может регулировать уровни объектных сигналов, сформированных посредством объектного декодера 121 согласно управляющей информации. Далее объектный декодер 121 может декодировать объектные сигналы, уровни которых отрегулированы. Соответственно, микшер/рендерер 123 не должен регулировать уровни декодированных объектных сигналов, предоставляемых посредством объектного декодера 121, а просто компонует декодированные объектные сигналы, предоставляемые посредством объектного декодера 121, в многоканальном пространстве. Вкратце, поскольку объектный декодер 121 регулирует уровни объектных сигналов, формируемых посредством объектного декодера 121, согласно управляющей информации микшер/рендерер 123 может легко компоновать объектные сигналы, формируемые посредством объектного декодера 121, в многоканальном пространстве без необходимости дополнительно регулировать уровни объектных сигналов, формируемых посредством объектного декодера 121. Следовательно, можно снижать сложность микширования/рендеринга.

Согласно варианту осуществления по фиг.2 объектный декодер устройства 120 декодирования аудио может адаптивно выполнять операцию декодирования посредством анализа управляющей информации, тем самым снижая сложность декодирования и сложность микширования/рендеринга. Может быть использована комбинация вышеописанных способов, выполняемых посредством устройства 120 декодирования аудио.

Фиг.3 иллюстрирует блок-схему устройства 130 декодирования аудио согласно второму варианту осуществления настоящего изобретения. Ссылаясь на фиг.3, устройство 130 декодирования аудио включает в себя объектный декодер 131 и микшер/рендерер 133. Устройство 130 декодирования аудио отличается за счет предоставления дополнительной информации не только в объектный декодер 131, но также в микшер/рендерер 133.

Устройство 130 декодирования аудио может эффективно выполнять операцию декодирования, даже когда имеется объектный сигнал, соответствующий периоду молчания. Например, второй-четвертый объектные сигналы могут соответствовать периоду воспроизведения музыки, в течение которого воспроизводится музыкальный инструмент, а первый объектный сигнал может соответствовать периоду приглушения звука, в течение которого воспроизводится только фоновая музыка, и первый объектный сигнал может соответствовать периоду молчания, в течение которого воспроизводится аккомпанемент. В этом случае информация, указывающая то, какой

из множества объектных сигналов соответствует периоду молчания, может быть включены в дополнительную информацию, и дополнительная информация может быть предоставлена в микшер/рендерер 133, а также в объектный декодер 131.

5 Объектный декодер 131 может минимизировать скорость декодирования не только посредством декодирования объектного сигнала, соответствующего периоду молчания. Объектный декодер 131 задает объектный сигнал, соответствующий значению в 0, и передает уровень объектного сигнала в микшер/рендерер 133. В  
10 общем, объектные сигналы, имеющие значение в 0, интерпретируются так же, как и объектные сигналы, имеющие значение, отличное от 0, и тем самым подвергаются операции микширования/рендеринга.

С другой стороны, устройство 130 декодирования аудио передает дополнительную информацию, включающую в себя информацию, указывающую то, какой из  
15 множества объектных сигналов соответствует периоду молчания, в микшер/рендерер 133, и тем самым не допускает обработки объектного сигнала, соответствующего периоду молчания, посредством операции микширования/рендеринга, выполняемой посредством микшера/рендерера 133. Следовательно, устройство 130 декодирования аудио позволяет не допускать лишнего  
20 возрастания сложности микширования/рендеринга.

Фиг.4 иллюстрирует блок-схему устройства 140 декодирования аудио согласно третьему варианту осуществления настоящего изобретения. Ссылаясь на фиг.4, устройство 140 декодирования аудио использует многоканальный декодер 141 вместо  
25 объектного декодера и микшера/рендерера и декодирует ряд объектных сигналов после того, как объектные сигналы надлежащим образом скомпонованы в многоканальном пространстве.

Более конкретно, устройство 140 декодирования аудио включает в себя многоканальный декодер 141 и преобразователь 145 параметров.  
30 Многоканальный 141 декодер формирует многоканальный сигнал, объектные сигналы которого уже скомпонованы в многоканальном пространстве, на основе сигнала понижающего микширования и информации пространственных параметров, которая является информацией основывающихся на каналах параметров, предоставляемой посредством преобразователя 145 параметров. Преобразователь 145  
35 параметров анализирует дополнительную информацию и управляющую информацию, передаваемую посредством устройства кодирования аудио (не показано), и формирует информацию пространственных параметров на основе результата анализа. Более конкретно, преобразователь 145 параметров формирует информацию  
40 пространственных параметров посредством комбинирования дополнительной информации и управляющей информации, которая включает в себя информацию настроек воспроизведения и информацию микширования. Т.е. преобразователь 145 параметров выполняет преобразование комбинации дополнительной информации и управляющей информации в пространственные данные, соответствующие модулю  
45 один-в-два (ОТТ) или модулю два-в-три (ТТТ).

Устройство 140 декодирования аудио может выполнять операцию многоканального декодирования, в которую объединены операция основывающегося на объектах декодирования и операция микширования/рендеринга, и тем самым  
50 может пропускать декодирование каждого объектного сигнала. Следовательно, можно снижать сложность декодирования и/или микширования/рендеринга.

Например, когда имеется 10 объектных сигналов, и многоканальный сигнал, полученный на основе 10 объектных сигналов, должен быть воспроизведен

5 посредством 5.1-канальной акустической системы, типичное устройство декодирования основывающегося на объектах аудио формирует декодированные сигналы, надлежащим образом соответствующие 10 объектным сигналам, на основе сигнала понижающего микширования и дополнительной информации, и затем формирует 5.1-канальный сигнал посредством надлежащей компоновки 10 объектных сигналов в многоканальное пространство, так что объектные сигналы могут стать подходящими для 5.1-канального акустического окружения. Тем не менее, недостаточно сформировать 10 объектных сигналов в ходе формирования 5.1-канального сигнала, и эта проблема становится более серьезной по мере того, как разность между числом объектных сигналов и числом каналов многоканального сигнала, который должен быть сформирован, возрастает.

15 С другой стороны, в варианте осуществления по фиг.4 устройство 140 декодирования аудио формирует информацию пространственных параметров, подходящую для 5.1-канального сигнала, на основе дополнительной информации и управляющей информации и предоставляет информацию пространственных параметров и сигнал понижающего микширования в многоканальный декодер 141. Затем многоканальный декодер 141 формирует 5.1-канальный сигнал на основе информации пространственных параметров и сигнала понижающего микширования. Другими словами, когда число каналов, которые должны быть выведены, составляет 5.1 каналов, устройство 140 декодирования аудио может просто сформировать 5.1-канальный сигнал на основе сигнала понижающего микширования без необходимости формировать 10 объектных сигналов и, таким образом, является более эффективным, чем традиционное устройство декодирования аудио, в отношении сложности.

30 Устройство 140 декодирования аудио считается эффективным, когда объем вычислений, требуемых для того, чтобы вычислять информацию пространственных параметров, соответствующую каждому из ОТТ-модуля и ТТТ-модуля посредством анализа дополнительной информации и управляющей информации, передаваемой посредством устройства кодирования аудио, меньше объема вычислений, требуемого для того, чтобы выполнять операцию микширования/рендеринга после декодирования каждого объектного сигнала.

35 Устройство 140 декодирования аудио может быть получено посредством добавления модуля для формирования информации пространственных параметров посредством анализа дополнительной информации и управляющей информации в типичное устройство декодирования многоканального аудио и поэтому может 40 сохранять совместимость с типичным устройством декодирования многоканального аудио. Так же, устройство 140 декодирования может повышать качество звука с использованием существующих средств типичного устройства декодирования многоканального аудио, таких как формирователь огибающей, средство временной обработки подполос (STP) и декоррелятор. С учетом всего этого следует сделать вывод о том, что все преимущества типичного способа декодирования 45 многоканального аудио могут быть легко применены к способу декодирования объектного аудио.

50 Информация пространственных параметров, передаваемая в многоканальный декодер 141 посредством преобразователя 145 параметров, может быть сжата, с тем, чтобы быть подходящей для передачи. Альтернативно, информация пространственных параметров может иметь такой же формат, что и формат данных, передаваемых посредством типичного устройства многоканального кодирования. Т.е.

информация пространственных параметров может быть подвергнута операции декодирования Хаффмана или операции контрольного декодирования и тем самым может быть передана в каждый модуль как несжатые данные пространственных меток. Первое подходит для передачи информации пространственных параметров в устройство декодирования многоканального аудио в удаленном месте, а второе удобно, поскольку нет необходимости устройству декодирования многоканального аудио преобразовывать сжатые данные пространственных меток в несжатые данные пространственных меток, которые могут быть легко использованы в операции декодирования.

Конфигурация информации пространственной задержки на основе анализа дополнительной информации и управляющей информации может вызывать задержку. Чтобы компенсировать эту задержку, может быть предусмотрен дополнительный буфер для сигнала понижающего микширования с тем, чтобы задержка между сигналом понижающего микширования и потоком битов могла компенсироваться. Альтернативно, может быть предусмотрен дополнительный буфер для информации пространственных параметров, полученной из управляющей информации с тем, чтобы задержка между информацией пространственных параметров и потоком битов могла компенсироваться. Эти способы, тем не менее, являются неудобными из-за необходимости предоставлять дополнительный буфер. Альтернативно, дополнительная информация может передаваться впереди сигнала понижающего микширования с учетом возможности возникновения задержки между сигналом понижающего микширования и информацией пространственных параметров. В этом случае информация пространственных параметров, полученная посредством комбинирования дополнительной информации и управляющей информации, не обязательно должна корректироваться, а может легко быть использована.

Если множество объектных сигналов из сигнала понижающего микширования имеют различные уровни, модуль произвольного усиления понижающего микширования (ADG), который может непосредственно компенсировать сигнал понижающего микширования, может определять относительные уровни объектных сигналов, и каждый из объектных сигналов может быть назначен предварительно определенной позиции в многоканальном пространстве с помощью данных пространственных меток, такие как информация разности уровней каналов (CLD), информация межканальных корреляций (ICC) и информация коэффициентов прогнозирования каналов (CPC).

Например, если управляющая информация указывает то, что предварительно определенный объектный сигнал должен быть назначен предварительно определенной позиции в многоканальном пространстве и имеет более высокий уровень, чем другие объектные сигналы, типичный многоканальный декодер может вычислять разность между энергиями каналов в сигнале понижающего микширования и поделить сигнал понижающего микширования на число выходных каналов на основе результатов вычислений. Тем не менее, типичный многоканальный декодер не может повышать или понижать громкость определенного звука в сигнале понижающего микширования. Другими словами, типичный многоканальный декодер просто распределяет сигнал понижающего микширования по числу выходных каналов и тем самым не может повышать или понижать громкость звука в сигнале понижающего микширования.

Относительно просто назначать каждый из ряда объектных сигналов в сигнале понижающего микширования, сформированном посредством объектного декодера,

предварительно определенной позиции в многоканальном пространстве согласно управляющей информации. Тем не менее, специальные методики требуются для того, чтобы увеличивать или уменьшать амплитуду предварительно определенного объектного сигнала. Другими словами, если сигнал понижающего микширования, сформированный посредством объектного декодера, используется как есть, трудно уменьшать амплитуду каждого объектного сигнала в сигнале понижающего микширования.

Следовательно, согласно варианту осуществления настоящего изобретения относительные амплитуды объектных сигналов могут варьироваться согласно управляющей информации посредством использования ADG-модуля 147, проиллюстрированного на фиг.5. ADG-модуль 147 может быть установлен в многоканальном декодере 141 или может быть отдельным от многоканального декодера 141.

Если относительные амплитуды объектных сигналов в сигнале понижающего микширования надлежащим образом отрегулированы с помощью ADG-модуля 147, можно выполнять объектное декодирование с помощью типичного многоканального декодера. Если сигнал понижающего микширования, сформированный посредством объектного декодера, является моно- или стереосигналом либо многоканальным сигналом с тремя или более каналами, сигнал понижающего микширования может быть обработан посредством ADG-модуля 147. Если сигнал понижающего микширования, сформированный посредством объектного декодера, имеет два или более каналов, и предварительно определенный объектный сигнал, который должен быть отрегулирован посредством ADG-модуля 147, существует только в одном из каналов сигнала понижающего микширования, ADG-модуль 147 может быть применен только к каналу, включающему в себя предварительно определенный объектный сигнал, вместо применения ко всем каналам сигнала понижающего микширования. Сигнал понижающего микширования, обработанный посредством ADG-модуля 147 вышеописанным способом, может быть легко обработан с помощью типичного многоканального кодера без необходимости модифицировать структуру многоканального декодера.

Даже когда конечный выходной сигнал не является многоканальным сигналом, который может быть воспроизведен посредством многоканальной акустической системы, а является стереофоническим (бинацеальным) сигналом, ADG-модуль 147 может быть использован для того, чтобы регулировать относительные амплитуды объектных сигналов конечного выходного сигнала.

Альтернативно применению ADG-модуля 147, информация усиления, задающая значение усиление, которое должно быть применено к каждому объектному сигналу, может быть включена в управляющую информацию в ходе формирования ряда объектных сигналов. Для этого структура типичного многоканального декодера может быть модифицирована. Несмотря на необходимость модификации структуры существующего многоканального декодера, этот способ является удобным в отношении сложности декодирования за счет применения значения усиления к каждому объектному сигналу в ходе операции декодирования без необходимости вычислять ADG и компенсировать каждый объектный сигнал.

ADG-модуль 147 может быть использован не только для регулирования объектных сигналов, но также для модификации информации спектра конкретного объектного сигнала. Более конкретно, ADG-модуль 147 может быть использован не только для того, чтобы повышать или понижать уровень конкретного объектного сигнала, но

также для того, чтобы модифицировать информацию спектра конкретного объектного сигнала, такую как усиление высокочастотной или низкочастотной части конкретного объектного сигнала. Невозможно модифицировать информацию спектра без использования ADG-модуля 147.

5 Фиг.6 иллюстрирует блок-схему устройства 150 декодирования аудио согласно четвертому варианту осуществления настоящего изобретения. Ссылаясь на фиг.6, устройство 150 декодирования аудио включает в себя многоканальный стереофонический декодер 151, первый преобразователь 157 параметров и второй преобразователь 159 параметров.

10 Второй преобразователь 159 параметров анализирует дополнительную информацию и управляющую информацию, которая предоставляется посредством устройства кодирования аудио, и конфигурирует информацию пространственных параметров на основе результата анализа. Первый преобразователь 157 параметров конфигурирует информацию виртуальных трехмерных (3D) параметров, которая может быть использована посредством многоканального стереофонического декодера 151, посредством добавления трехмерной (3D) информации, такой как функция моделирования восприятия звука (HRTF), в информацию пространственных параметров. Многоканальный стереофонический декодер 151 формирует стереофонический сигнал посредством применения информации стереофонических параметров к сигналу понижающего микширования.

20 Первый преобразователь 157 параметров и второй преобразователь 159 параметров могут быть заменены посредством одного модуля, т.е. модуля 155 преобразования параметров, который принимает дополнительную информацию, управляющую информацию и 3D информацию и конфигурирует информацию стереофонических параметров на основе дополнительной информации, управляющей информацией и HRTF-параметров.

30 Традиционно для того, чтобы сформировать стереофонический сигнал для воспроизведения сигнала понижающего микширования, включающего в себя 10 объектных сигналов, с помощью наушников, объектный сигнал должен сформировать 10 декодированных сигналов, надлежащим образом соответствующих 10 объектным сигналам на основе сигнала понижающего микширования и дополнительной информации. Затем микшер/рендерер назначает каждый из 10 объектных сигналов предварительно определенной позиции в многоканальном пространстве со ссылкой на управляющую информацию, с тем, чтобы удовлетворять требованиям 5-канального акустического окружения. После этого микшер/рендерер формирует 5-канальный сигнал, который может быть воспроизведен посредством 5-канальной акустической системы. Далее микшер/рендерер применяет 3D информацию к 5-канальному сигналу, тем самым формируя 2-канальный сигнал. Вкратце, вышеупомянутый традиционный способ декодирования аудио включает в себя воспроизведение 10 объектных сигналов, преобразование 10 объектных сигналов в 5-канальный сигнал и формирование 2-канального сигнала на основе 5-канального сигнала, и это тем самым является неэффективным.

45 С другой стороны, устройство 150 декодирования аудио может легко формировать стереофонический сигнал, который может быть воспроизведен с помощью наушников, на основе объектных сигналов. Помимо этого устройство 150 декодирования аудио конфигурирует информацию пространственных параметров посредством анализа дополнительной информации и управляющей информацией и тем

самым может формировать стереофонический сигнал с помощью типичного многоканального стереофонического декодера. Более того, устройство 150 декодирования аудио может использовать типичный многоканальный стереофонический декодер, даже когда оснащено встроенным преобразователем параметров, который принимает дополнительную информацию, управляющую информацией и HRTF-параметры и конфигурирует информацию стереофонических параметров на основе дополнительной информации, управляющей информацией и HRTF-параметров.

Фиг.7 иллюстрирует блок-схему устройства 160 декодирования аудио согласно пятому варианту осуществления настоящего изобретения. Ссылаясь на фиг.7, устройство 160 декодирования аудио включает в себя препроцессор 161, многоканальный 163 декодер и преобразователь 165 параметров.

Преобразователь 165 параметров формирует информацию пространственных параметров, которая может быть использована посредством многоканального декодера 163, и информацию параметров, которая может быть использована посредством препроцессора 161. Препроцессор 161 выполняет операцию предварительной обработки с сигналом понижающего микширования и передает сигнал понижающего микширования, получающийся в результате операции предварительной обработки, в многоканальный декодер 163. Многоканальный декодер 163 выполняет операцию декодирования сигнала понижающего микширования, передаваемого посредством препроцессора 161, тем самым выводя стереосигнал, стереофонический стереосигнал и многоканальный сигнал. Примеры операции предварительной обработки, выполняемой посредством препроцессора 161, включают в себя модификацию или преобразование сигнала понижающего микширования во временной области или частотной области с помощью фильтрации.

Если сигнал понижающего микширования, вводимый в устройство 160 декодирования аудио, является стереосигналом, сигнал понижающего микширования, возможно, должен быть подвергнут обработке понижающего микширования, выполняемой посредством препроцессора 161, перед вводом в многоканальный декодер 163, поскольку многоканальный декодер 163 не может преобразовывать объектный сигнал, соответствующий левому каналу стереосигнала понижающего микширования, в правый канал многоканального сигнала посредством декодирования. Следовательно, чтобы сдвигать объектный сигнал, принадлежащий левому каналу стереосигнала понижающего микширования, к правому каналу, стереосигнал понижающего микширования, возможно, должен быть предварительно обработан посредством препроцессора 161, и предварительно обработанный сигнал понижающего микширования может быть введен в многоканальный декодер 163.

Предварительная обработка стереосигнала понижающего микширования может выполняться на основе информации предварительной обработки, полученной из дополнительной информации и из управляющей информации.

Фиг.8 иллюстрирует блок-схему устройства 170 декодирования аудио согласно шестому варианту осуществления настоящего изобретения. Ссылаясь на фиг.8, устройство 170 декодирования аудио включает в себя многоканальный 171 декодер, постпроцессор 173 и преобразователь 175 параметров.

Преобразователь 175 параметров формирует информацию пространственных параметров, которая может быть использована посредством многоканального декодера 163, и информацию параметров, которая может быть использована посредством постпроцессора 173. Постпроцессор 173 выполняет операцию

постобработки с сигналом, выводимым посредством многоканального декодера 173. Примеры сигнала, выводимого посредством многоканального декодера 173, включают в себя стереосигнал, стереофонический стереосигнал и многоканальный сигнал.

5       Примеры операции постобработки, выполняемой посредством постпроцессора 173, включают в себя модификацию и преобразование каждого канала или всех каналов выходного сигнала. Например, если дополнительная информация включает в себя  
10       информацию собственной частоты, касающуюся предварительно определенного объектного сигнала, постпроцессор 173 может удалять гармонические компоненты из предварительно определенного объектного сигнала со ссылкой на информацию  
15       собственной частоты. Способ декодирования многоканального аудио может быть недостаточно эффективным для того, чтобы использовать в системе караоке. Тем не менее, если информация собственной частоты, касающаяся вокальных объектных  
20       сигналов, включена в дополнительную информацию, и гармонические компоненты вокальных объектных сигналов удаляются в ходе операции постобработки, можно реализовывать высокопроизводительную систему караоке посредством использования  
25       варианта осуществления по фиг.8. Вариант осуществления по фиг.8 также может быть применен к объектным сигналам, отличным от вокальных объектных сигналов. Например, можно удалять звук предварительно определенного музыкального  
30       инструмента посредством использования варианта осуществления по фиг.8. Так же, можно усиливать предварительно определенные гармонические компоненты с помощью информации собственной частоты, касающейся объектных сигналов,  
35       посредством использования варианта осуществления по фиг.8. Вкратце, параметры постобработки могут предоставлять возможность применения различных эффектов, таких как вставка эффекта реверберации, добавление шума и усиление низкой части, которые не могут быть применены посредством многоканального декодера 171.

40       Постпроцессор 173 может непосредственно применять дополнительный эффект к сигналу понижающего микширования или добавлять сигнал понижающего микширования, к которому эффект уже применен, в вывод многоканального декодера 171. Постпроцессор 173 всегда может изменять спектр объекта или модифицировать сигнал понижающего микширования при необходимости. Если не  
45       подходит непосредственно выполнять операцию обработки эффектов, такую как реверберация, для сигнала понижающего микширования и передавать сигнал, полученный посредством операции обработки эффектов, в многоканальный декодер 171, препроцессор 173 может просто добавить сигнал, полученный  
50       посредством операции обработки эффектов, в вывод многоканального декодера 171 вместо непосредственного выполнения обработки эффектов с сигналом понижающего микширования и передачи результата обработки эффектов в многоканальный декодер 171.

55       Фиг.9 иллюстрирует блок-схему устройства 180 декодирования аудио согласно седьмому варианту осуществления настоящего изобретения. Ссылаясь на фиг.9, устройство 180 декодирования аудио включает в себя препроцессор 181, многоканальный 183 декодер, постпроцессор 185 и преобразователь 187 параметров.

60       Описание препроцессора 161 непосредственно применимо к препроцессору 181. Постпроцессор 185 может быть использован для того, чтобы суммировать вывод препроцессора 181 и вывод многоканального декодера 185 и тем самым предоставлять конечный сигнал. В этом случае постпроцессор 185 выступает просто в качестве сумматора для суммирования сигналов. Параметр эффектов может

предоставляться в тот из препроцессора 181 и постпроцессора 185, который осуществляет применение эффекта. Помимо этого, добавление сигнала, полученного посредством применения эффекта сигнала понижающего микширования, в вывод многоканального декодера 183 и применение эффекта к выводу многоканального декодера 185 может выполняться одновременно.

Препроцессоры 161 и 181 по фиг.7 и 9 могут выполнять рендеринг сигнала понижающего микширования согласно управляющей информации, предоставленной пользователем. Помимо этого, препроцессоры 161 и 181 по фиг.7 и 9 могут повышать или понижать уровни объектных сигналов и изменять спектр объектных сигналов. В этом случае препроцессоры 161 и 181 по фиг.7 и 9 могут выполнять функции ADG-модуля.

Рендеринг объектного сигнала согласно информации направления объектного сигнала, регулировка уровня объектного сигнала и изменение спектра объектного сигнала могут выполняться одновременно. Помимо этого часть рендеринга объектного сигнала согласно информации направления объектного сигнала, регулировки уровня объектного сигнала и изменения спектра объектного сигнала может выполняться с помощью препроцессора 161 или 181, а то из рендеринга объектного сигнала согласно информации направления объектного сигнала, регулировки уровня объектного сигнала и изменения спектра объектного сигнала, что не выполняется посредством препроцессора 161 или 181, может выполняться посредством ADG-модуля. Например, неэффективно изменять спектр объектного сигнала посредством использования ADG-модуля, который использует интервал уровня квантования и интервал параметрической полосы. В этом случае препроцессор 161 или 181 может быть использован для того, чтобы поминутно изменять спектр объектного сигнала на почастотной основе, а ADG-модуль может быть использован для того, чтобы регулировать уровень объектного сигнала.

Фиг.10 иллюстрирует блок-схему устройства декодирования аудио согласно восьмому варианту осуществления настоящего изобретения. Ссылаясь на фиг.10, устройство 200 декодирования аудио включает в себя формирователь 201 матриц рендеринга, транскодер 203, многоканальный декодер 205, препроцессор 207, процессор 208 эффектов и сумматор 209.

Формирователь 201 матриц рендеринга формирует матрицу рендеринга, которая представляет информацию позиций объектов, касающуюся позиций объектных сигналов, и информацию конфигурации воспроизведения, касающуюся уровней объектных сигналов, и предоставляет матрицу рендеринга в транскодер 203. Формирователь 201 матриц рендеринга формирует 3D информацию, такую как HRTF-коэффициент, на основе информации позиции объектов. HRTF - это передаточная функция, которая описывает передачу звуковых волн между источником звука в произвольной позиции и барабанной перепонкой, и возвращает значение, которое варьируется согласно направлению и высоте источника звука. Если сигнал без направленности фильтруется с помощью HRTF, сигнал может быть услышан, как если бы он воспроизводился из определенного направления.

Информация позиции объектов и информация конфигурации воспроизведения, которая принимает посредством формирователя 201 матриц рендеринга, может варьироваться во времени и может предоставляться конечным пользователем.

Транскодер 203 формирует основывающуюся на каналах дополнительную информацию на основе основывающейся на объектах дополнительной информации, матрицы рендеринга и 3D информации, и предоставляет в многоканальный

декодер 209 основывающуюся на каналах дополнительную информацию и 3D информацию, требуемую для многоканального декодера 209. Т.е. транскодер 203 передает основывающуюся на каналах дополнительную информацию, касающуюся M каналов, которая получается из информации основывающихся на объектах параметров, касающейся N объектных сигналов, и 3D информации по каждому из N объектных сигналов, в многоканальный декодер 205.

Многоканальный декодер 205 формирует многоканальный аудиосигнал на основе сигнала понижающего микширования и основывающейся на каналах дополнительной информации, предоставленной посредством транскодера 203, и выполняет 3D рендеринг многоканального аудиосигнала согласно 3D информации, тем самым формируя 3D многоканальный сигнал. Формирователь 201 матриц рендеринга может включать в себя базу данных 3D информации (не показана).

Если есть необходимость предварительно обрабатывать сигнал понижающего микширования перед вводом сигнала понижающего микширования в многоканальный декодер 205, транскодер 203 передает информацию, касающуюся предварительной обработки, в препроцессор 207. Основывающаяся на объектах дополнительная информация включает в себя информацию, касающуюся всех объектных сигналов, а матрица рендеринга включает в себя информацию позиции объектов и информацию конфигурации воспроизведения. Транскодер 203 может формировать основывающуюся на каналах дополнительную информацию на основе основывающейся на объектах дополнительной информации и матрицы рендеринга, и затем формирует основывающуюся на каналах дополнительную информацию, требуемую для микширования и воспроизведения объектных сигналов согласно канальной информации. Затем транскодер 203 передает основывающуюся на каналах дополнительную информацию в многоканальный декодер 205.

Основывающаяся на каналах дополнительная информация и 3D информация, предоставляемая посредством транскодера 203, может включать в себя индексы кадров. Таким образом, многоканальный декодер 205 может синхронизировать основывающуюся на каналах дополнительную информацию и 3D информацию посредством использования индексов кадров, и тем самым может иметь возможность применять 3D информацию только к определенным кадрам потока битов. Помимо этого, даже если 3D информация обновлена, можно легко синхронизировать основывающуюся на каналах дополнительную информацию и обновленную 3D информацию посредством использования индексов кадров. Т.е. индексы кадров могут быть включены в основывающуюся на каналах дополнительную информацию и 3D информацию, соответственно, чтобы многоканальный декодер 205 синхронизировал основывающуюся на каналах дополнительную информацию и 3D информацию.

Препроцессор 207 может выполнять предварительную обработку входного сигнала понижающего микширования, если необходимо, перед тем как входной сигнал понижающего микширования вводится в многоканальный декодер 205. Как описано выше, если входной сигнал понижающего микширования является стереосигналом, и нет необходимости воспроизводить объектный сигнал, принадлежащий левому каналу, из правого канала, сигнал понижающего микширования, возможно, должен быть подвергнут предварительной обработке, выполняемой посредством препроцессора 207, перед вводом в многоканальный декодер 205, поскольку многоканальный декодер 205 не может сдвигать объектный сигнал из одного канала к другому. Информация, требуемая для предварительной обработки входного сигнала понижающего микширования, может предоставляться в препроцессор 207

посредством транскодера 205. Сигнал понижающего микширования, получаемый посредством предварительной обработки, выполняемой посредством препроцессора 207, может передаваться в многоканальный декодер 205.

5 Процессор 208 эффектов и сумматор 209 могут непосредственно применять дополнительный эффект к сигналу понижающего микширования или добавлять сигнал понижающего микширования, к которому эффект уже применен, в вывод многоканального декодера 205. Процессор 208 эффектов всегда может изменять спектр объекта или модифицировать сигнал понижающего микширования при  
10 необходимости. Если не подходит непосредственно выполнять операцию обработки эффектов, такую как реверберация, для сигнала понижающего микширования и передавать сигнал, полученный посредством операции обработки эффектов, в многоканальный декодер 205, процессор 208 эффектов может просто добавить сигнал, полученный посредством операции обработки эффектов, в вывод многоканального  
15 декодера 205 вместо непосредственного выполнения обработки эффектов с сигналом понижающего микширования и передачи результата обработки эффектов в многоканальный декодер 205.

20 Далее подробно описывается матрица рендеринга, формируемая посредством формирователя 201 матриц рендеринга.

Матрица рендеринга - это матрица, которая представляет позиции и конфигурацию воспроизведения объектных сигналов. Т.е. если имеется  $N$  объектных сигналов и  $M$  каналов, матрица рендеринга может указывать то, как  $N$  объектных сигналов преобразуются в  $M$  каналов различными способами.

25 Более конкретно, когда  $N$  объектных сигналов отображается на  $M$  каналов, матрица рендеринга  $N \times M$  может быть установлена. В этом случае матрица рендеринга включает в себя  $N$  строк, которые, соответственно, представляют  $N$  объектных сигналов, и  $M$  столбцов, которые, соответственно, представляют  $M$  каналов. Каждый из  $M$  коэффициентов в каждой из  $N$  строк может быть действительным числом или  
30 целым числом, указывающим отношение части объектного сигнала, назначенной соответствующему каналу, ко всему объектному сигналу.

Более конкретно,  $M$  коэффициентов в каждой из  $N$  строк матрицы рендеринга  $N \times M$  могут быть действительными числами. В таком случае, если сумма  $M$  коэффициентов в строке матрицы рендеринга  $N \times M$  равна предварительно определенному опорному значению, например 1, может быть определено то, что уровень объектного сигнала не варьировался. Если сумма  $M$  коэффициентов меньше 1, определяется то, что уровень объектного сигнала понизился. Если сумма  $M$  коэффициентов больше 1, определяется то, что уровень объектного сигнала повысился. Предварительно определенным опорным значением может быть числовое значение, отличное от 1. Величина, на которую варьируется уровень объектного сигнала, может быть ограничена диапазоном в 12 дБ. Например, если предварительно определенное опорное значение равно 1, а сумма  $M$  коэффициентов составляет 1,5, может быть определено то, что  
40 уровень объектного сигнала увеличился на 12 дБ. Если предварительно определенное опорное значение равно 1, а сумма  $M$  коэффициентов составляет 0,5, определяется то, что уровень объектного сигнала уменьшился на 12 дБ. Если предварительно определенное опорное значение равно 1, а сумма  $M$  коэффициентов составляет 0,5-1,5, определяется то, что объектный сигнал варьируется на предварительно определенную величину между -12 дБ и +12 дБ, и предварительно определенная величина может быть линейно определена согласно сумме из  $M$  коэффициентов.

50  $M$  коэффициентов в каждой из  $N$  строк матрицы рендеринга  $N \times M$  могут быть

целыми числами. В таком случае, если сумма  $M$  коэффициентов в строке матрицы рендеринга  $N \times M$  равна предварительно определенному опорному значению, например, 10, 20, 30 или 100, может быть определено то, что уровень объектного сигнала не варьировался. Если сумма  $M$  коэффициентов меньше предварительно  
 5 определенного опорного значения, может быть определено то, что уровень объектного сигнала не понизился. Если сумма  $M$  коэффициентов больше предварительно определенного опорного значения, может быть определено то, что уровень объектного сигнала не повысился. Величина, на которую варьируется  
 10 уровень объектного сигнала, может быть ограничена, например, диапазоном в 12 дБ. Величина, на которую сумма  $M$  коэффициентов отличается от предварительно заданного опорного значения, может представлять величину (в дБ), на которую варьировался уровень объектного сигнала. Например, если сумма  $M$  коэффициентов  
 15 больше предварительно определенного опорного значения, может быть определено то, что уровень объектного сигнала повысился на 2 дБ. Следовательно, если предварительно определенное опорное значение равно 20, а сумма  $M$  коэффициентов составляет 23, может быть определено то, что уровень объектного сигнала увеличился на 6 дБ. Если предварительно определенное опорное значение равно 20, а сумма  $M$   
 20 коэффициентов составляет 15, может быть определено то, что уровень объектного сигнала уменьшился на 10 дБ.

Например, если имеется шесть объектных сигналов и пять каналов (т.е. передний левый (FL), передний правый (FR), центральный (C), задний левый (RL) и задний  
 25 правый (RR) каналы), матрица рендеринга  $6 \times 5$ , имеющая шесть строк, надлежащим образом соответствующих шести объектным сигналам, и пять столбцов, надлежащим образом соответствующих пяти каналам, может быть установлена. Коэффициенты матрицы рендеринга  $6 \times 5$  могут быть целыми числами, указывающими соотношение, с которым каждый из шести объектных сигналов распределяется по пяти каналам.  
 30 Матрица рендеринга  $6 \times 5$  может иметь опорное значение в 10. Таким образом, если сумма пяти коэффициентов в любой одной из шести строк матрицы рендеринга  $6 \times 5$  равна 10, может быть определено, что уровень соответствующего сигнала варьировался. Величина, на которую сумма пяти коэффициентов в любой одной из  
 35 шести строк матрицы рендеринга  $6 \times 5$  отличается от опорного значения, представляет величину, на которую варьировался уровень соответствующего объектного сигнала. Например, если сумма пяти коэффициентов в любой одной из шести строк матрицы рендеринга  $6 \times 5$  отличается от опорного значения на 1, может быть определено, что уровень соответствующего объектного сигнала варьировался на 2 дБ. Матрица  
 40 рендеринга  $6 \times 5$  может быть представлена посредством уравнения (1):

Уравнение 1

$$\begin{bmatrix}
 3 & 1 & 2 & 2 & 2 \\
 2 & 4 & 3 & 1 & 2 \\
 45 & 0 & 0 & 12 & 0 & 0 \\
 7 & 0 & 0 & 0 & 0 \\
 2 & 2 & 2 & 2 & 2 \\
 2 & 1 & 1 & 2 & 1
 \end{bmatrix}$$

50 Ссылаясь на матрицу рендеринга  $6 \times 5$  по уравнению (1), первая строка соответствует первому объектному сигналу и представляет соотношение, с которым первый объектный сигнал распределяется между FL, FR, C, RL и RR каналами.

Поскольку первый коэффициент первой строки имеет наибольшее целое значение в 3, и сумма коэффициентов первой строки равна 10, определяется то, что первый объектный сигнал главным образом распределяется в FL канал и что уровень первого объектного сигнала не варьировался. Поскольку второй коэффициент второй строки, которая соответствует второму объектному сигналу, имеет наибольшее целое значение в 4 и сумма коэффициентов второй строки равна 12, определяется то, что второй объектный сигнал главным образом распределяется в FR канал и что уровень второго объектного сигнала увеличился на 4 дБ. Поскольку третий коэффициент третьей строки, которая соответствует третьему объектному сигналу, имеет наибольшее целое значение в 12 и сумма коэффициентов третьей строки равна 12, определяется то, что третий объектный сигнал распределяется только в С канал и что уровень третьего объектного сигнала увеличился на 4 дБ. Поскольку все коэффициенты пятой строки, которая соответствует пятому объектному сигналу, имеют одинаковое целое значение в 2 и сумма коэффициентов пятой строки равна 10, определяется то, что пятый объектный сигнал одинаково распределяется между FL, FR, С, RL и RR каналами и что уровень пятого объектного сигнала не варьировался.

Альтернативно, когда  $N$  объектных сигналов отображаются на  $M$  каналов, матрица рендеринга  $N \times (M+1)$  может быть установлена. Матрица рендеринга  $N \times (M+1)$  очень похожа на матрицу рендеринга  $N \times M$ . Более конкретно, в матрице рендеринга  $N \times (M+1)$ , как в матрице рендеринга  $N \times M$ , с первый по  $M$ -й коэффициенты в каждой из  $N$  строк представляют соотношение, с которым соответствующий объектный сигнал распределяется между FL, FR, С, RL и RR каналами. Тем не менее, матрица рендеринга  $N \times (M+1)$ , в отличие от матрицы рендеринга  $N \times M$ , имеет дополнительный столбец (т.е.  $(M+1)$ -й столбец) для представления уровней объектных сигналов.

Матрица рендеринга  $N \times (M+1)$ , в отличие от матрицы рендеринга  $N \times M$ , указывает то, как объектный сигнал распределен между  $M$  каналами, а также то, варьировался ли уровень объектного сигнала по отдельности. Таким образом, посредством использования матрицы рендеринга  $N \times (M+1)$  можно легко получать информацию, касающуюся варьирования, если имеется, уровня объектного сигнала без необходимости дополнительных вычислений. Поскольку матрица рендеринга  $N \times (M+1)$  является практически такой же, что и матрица рендеринга  $N \times M$ , матрица рендеринга  $N \times (M+1)$  может быть легко преобразована в матрицу рендеринга  $N \times M$ , и наоборот, без необходимости дополнительной информации.

Еще альтернативно, когда  $N$  объектных сигналов отображается на  $M$  каналов, может быть установлена матрица рендеринга  $N \times 2$ . Матрица рендеринга  $N \times 2$  имеет первый столбец, указывающий угловые позиции объектных сигналов, и второй столбец, указывающий варьирование, если имеется, уровня каждого из объектных сигналов. Матрица рендеринга  $N \times 2$  может представлять угловые позиции объектных сигналов с регулярными интервалами в 1 или 3 градуса в пределах диапазона 0-360 градусов. Объектный сигнал, который равномерно распределен между всеми направлениями, может быть представлен посредством предварительно заданного значения, а не угла.

Матрица рендеринга  $N \times 2$  может быть преобразована в матрицу рендеринга  $N \times 3$ , которая может указывать не только 2D направления объектных сигналов, но также 3D направления объектных сигналов. Более конкретно, второй столбец матрицы рендеринга  $N \times 3$  может быть использован для того, чтобы указывать 3D направления объектных сигналов. Третий столбец матрицы рендеринга  $N \times 3$  указывает варьирование, если есть, уровня каждого объектного сигнала с помощью того же

способа, что и использован посредством матрицы рендеринга  $N \times M$ . Если конечный режим воспроизведения объектного декодера - это стереофоническое стерео, формирователь 201 матриц рендеринга может передавать 3D информацию, указывающую позицию каждого объектного сигнала, или индекс, соответствующий 3D информации. Во втором случае транскодеру 203 может потребоваться иметь 3D информацию, соответствующую индексу, переданному посредством формирователя 201 матриц рендеринга. Помимо того, если 3D информация, указывающая позицию каждого объектного сигнала, принимается от формирователя 201 матриц рендеринга, транскодер 203 может иметь возможность вычислять 3D информацию, которая может быть использована посредством многоканального кодера 205, на основе принятой 3D информации, матрицы рендеринга и основывающейся на объектах дополнительной информации.

Матрица рендеринга и 3D информация могут адаптивно варьироваться в реальном времени согласно модификации, выполненной в информацию позиции объектов и информацию конфигурации воспроизведения посредством конечного пользователя. Следовательно, информация, касающаяся того, обновляется ли матрица рендеринга и 3D информация, и обновления, если имеют место, в матрице рендеринга и 3D информации могут передаваться в транскодер 203 с регулярными временными интервалами, например, с интервалами в 0,5 с. Далее, если обновления в матрице рендеринга и 3D информации обнаружены, транскодер 203 может выполнить линейное преобразование принимаемых обновлений и существующей матрицы рендеринга и существующей 3D информации, при условии, что матрица рендеринга и 3D информация линейно варьируются во времени.

Если информация позиций объектов и информация конфигурации воспроизведения не модифицировалась конечным пользователем с момента передачи матрицы рендеринга и 3D информации в транскодер 203, информация, указывающая то, что матрица рендеринга и 3D информация не варьировалась, может быть передана в транскодер 203. С другой стороны, если информация позиций объектов и информация конфигурации воспроизведения модифицировалась посредством конечного пользователя с момента передачи матрицы рендеринга и 3D информации в транскодер 203, информация, указывающая то, что матрица рендеринга и 3D информация варьировалась, и обновления в матрице рендеринга и 3D информации могут быть переданы в транскодер 203. Более конкретно, обновления в матрице рендеринга и обновления 3D информации могут отдельно передаваться в транскодер 203. Альтернативно, обновления в матрице рендеринга и/или обновления 3D информации могут совместно представляться посредством предварительно заданного репрезентативного значения. Далее, предварительно заданное репрезентативное значение может передаваться в транскодер 203 вместе с информацией, указывающей то, что предварительно заданное репрезентативное значение соответствует обновлениям в матрице рендеринга или обновлениям 3D информации. Таким образом, можно легко оповещать транскодер 203 о том, обновлялась или нет матрица рендеринга и 3D информация.

Матрица рендеринга  $N \times M$ , аналогично указанной посредством уравнения (1), также может включать в себя дополнительный столбец для представления информации 3D направления объектных сигналов. В этом случае дополнительный столбец может представлять информацию 3D направления объектных сигналов как углы в диапазоне от  $-90$  до  $+90$  градусов. Дополнительный столбец может быть предусмотрен не только в матрице  $N \times M$ , но также в матрице рендеринга  $N \times (M+1)$  и в матрице  $N \times 2$ .

Информация 3D направления объектных сигналов может быть необязательной для использования в обычном режиме декодирования многоканального декодера. Вместо этого, информация 3D направления объектных сигналов может быть обязательной для использования в стереофоническом режиме многоканального декодера.

5 Информация 3D направления объектных сигналов может передаваться вместе с матрицей рендеринга. Альтернативно, информация 3D направления объектных сигналов может передаваться вместе с 3D информацией. Информация 3D направления объектных сигналов не влияет на основывающуюся на каналах дополнительную  
10 информацию, но влияет на 3D информацию при операции декодирования в стереофоническом режиме.

Информация, касающаяся пространственных позиций и уровней объектных сигналов, может предоставляться как матрица рендеринга. Альтернативно,  
15 информация, касающаяся пространственных позиций и уровней объектных сигналов, может представляться как модификации спектра объектного сигнала, такие как интенсификация низкочастотных частей или высокочастотных частей объектных сигналов. В этом случае информация, касающаяся модификаций спектра объектных сигналов, может передаваться как варьирования уровня в каждой параметрической  
20 полосе, которая используется в многоканальном кодеке. Если конечный пользователь контролирует модификации спектра объектных сигналов, информация, касающаяся модификаций спектра объектных сигналов, может передаваться как матрица спектра отдельно от матрицы рендеринга. Матрица спектра может иметь столько строк, сколько объектных сигналов, и столько столбцов, сколько параметров. Каждый  
25 коэффициент в матрице спектра указывает информацию, касающуюся регулирования уровня каждой параметрической полосы.

Далее подробно описывается работа транскодера 203. Транскодер 203 формирует основывающуюся на каналах дополнительную информацию для многоканального  
30 декодера 205 на основе основывающейся на объектах дополнительной информации, информации матриц рендеринга и 3D информации и передает основывающуюся на каналах дополнительную информацию в многоканальный декодер 205. Кроме того, транскодер 203 формирует 3D информацию для многоканального декодера 205 и передает 3D информацию в многоканальный декодер 205. Если входной сигнал  
35 понижающего микширования должен быть предварительно обработан перед вводом в многоканальный декодер 205, транскодер 203 может передавать информацию, касающуюся входного сигнала понижающего микширования.

Транскодер 203 может принимать основывающуюся на объектах дополнительную  
40 информацию, указывающую то, как множество объектных сигналов включены во входной сигнал понижающего микширования. Основывающаяся на объектах дополнительная информация может указывать то, как множество объектных сигналов включены во входной сигнал понижающего микширования, посредством использования ОТТ-модуля и ТТТ-модуля и посредством использования  
45 информации CLD, ICC и CPC. Основывающаяся на объектах дополнительная информация может предоставлять описания различных способов, которые могут выполняться посредством объектного кодера для указания информации, касающейся каждого из множества объектных сигналов, и тем самым может иметь возможность  
50 указывать то, как объектные сигналы включены в дополнительную информацию.

В случае ТТТ-модуля многоканального кодера L-, C- и R-сигналы могут микшироваться с понижением или микшироваться с повышением в L- и R-сигналы. В этом случае C-сигнал может совместно использовать младший бит L- и R-сигналов.

Тем не менее, это редко происходит в случае понижающего микширования или повышающего микширования объектных сигналов. Следовательно, ОТТ-модуль широко используется для того, чтобы выполнять повышающее микширование или понижающее микширование для объектного кодирования. Даже если С-сигнал  
5 включает в себя независимый компонент сигнала, а не части L- и R-сигналов, ТТТ-модуль может быть использован для того, чтобы выполнять повышающее микширование или понижающее микширование для объектного кодирования.

Например, если имеется шесть объектных сигналов, шесть объектных сигналов  
10 могут быть преобразованы в сигнал понижающего микширования посредством ОТТ-модуля, и информация, касающаяся каждого из объектных сигналов, может быть получена с помощью ОТТ-модуля, как проиллюстрировано на фиг.11.

Ссылаясь на фиг.11, шесть объектных сигналов могут быть представлены  
15 посредством одного сигнала понижающего микширования и информации (такой как информация CLD и ICC), предоставленной посредством всего пяти ОТТ-модулей 211, 213, 215, 217 и 219. Структура, проиллюстрированная на фиг.11, может быть изменена различными способами. Т.е. ссылаясь на фиг.11, первый ОТТ-модуль 211 может принимать два из шести объектных сигналов. Помимо того, способ, которым ОТТ-  
20 модули 211, 213, 215, 217 и 219 иерархически соединены, может свободно варьироваться. Следовательно, дополнительная информация может включать в себя информацию иерархической структуры, указывающую то, как ОТТ-модули 211, 213, 215, 217 и 219 иерархически соединены, и информацию входных позиций,  
25 указывающую то, в какой ОТТ-модуль каждый объектный сигнал вводится. Если ОТТ-модули 211, 213, 215, 217 и 219 формируют произвольную древовидную структуру, способ, используемый в многоканальном кодеке для представления произвольной древовидной структуры, может быть использован для того, чтобы указывать эту  
30 информацию иерархической структуры. Помимо того, такая информация входной позиции может быть указана различными способами.

Дополнительная информация также может включать в себя информацию, касающуюся периода приглушения звука каждого объектного сигнала. В этом случае древовидная структура ОТТ-модулей 211, 213, 215, 217 и 219 может адаптивно  
35 варьироваться во времени. Например, ссылаясь на фиг.11, когда первый объектный сигнал ОБЪЕСТ1 приглушен, информация, касающаяся первого ОТТ-модуля 211, является лишней, и только второй объектный сигнал ОБЪЕСТ2 может быть введен в четвертый ОТТ-модуль 217. Затем древовидная структура ОТТ-модулей 211, 213, 215, 217 и 219 может варьироваться соответствующим образом. Таким образом,  
40 информация, касающаяся варьирования, если оно имеет место, в древовидной структуре ОТТ-модулей 211, 213, 215, 217 и 219 может быть включена в дополнительную информацию.

Если предварительно определенный объектный сигнал приглушен, информация, указывающая то, что ОТТ-модуль, соответствующий предварительно определенному  
45 объектному сигналу, не используется, и информация, указывающая то, что нет доступных меток из ОТТ-модуля, может быть предоставлена. Таким образом, можно снижать размер дополнительной информации за счет невключения информации, касающейся ОТТ-модулей или ТТТ-модулей, которые не используются в  
50 дополнительной информации. Даже если древовидная структура множества ОТТ-или ТТТ-модулей модифицирована, можно легко определять то, какие из ОТТ-или ТТТ-модулей включены или выключены, на основе информации, указывающей то, что объектные сигналы приглушены. Следовательно, нет необходимости часто

передавать информацию, касающуюся модификаций, если имеются, в древовидной структуре ОТТ- или ТТТ-модулей. Вместо этого передается информация, указывающая то, что объектный сигнал приглушен. Затем декодер может легко определять то, какая часть древовидной структуры ОТТ- или ТТТ-модулей должна быть модифицирована.

Следовательно, можно минимизировать размер информации, которая должна быть передана в декодер. Помимо того, можно легко передавать метки, касающиеся объектных сигналов, в декодер.

Фиг.12 иллюстрирует схему для пояснения того, как множество объектных сигналов включаются в сигнал понижающего микширования. В варианте осуществления по фиг.11 структура ОТТ-модуля многоканального кодирования приспособляется как есть. Тем не менее, в варианте осуществления по фиг.12 используется вариант структуры ОТТ-модуля многоканального кодирования. Ссылаясь на фиг.12, множество объектных сигналов вводятся в каждый модуль, и только один сигнал понижающего микширования формируется в конце. Ссылаясь на фиг.12, информация, касающаяся каждого из множества объектных сигналов, может быть представлена посредством отношения уровня энергии каждого из объектных сигналов к общему уровню энергии объектных сигналов. Тем не менее, по мере того как число объектных сигналов возрастает, отношение уровня энергии каждого из объектных сигналов к общему уровню энергии объектных сигналов уменьшается. Чтобы разрешить это, находится один из множества объектных сигналов (далее упоминаемый как объектный сигнал с наибольшей энергией), имеющий наибольший уровень энергии в предварительно определенной параметрической полосе, и отношения уровней энергии других объектных сигналов (далее упоминаемых как объектные сигналы не с наибольшей энергией) к уровню энергии объектного сигнала с наибольшей энергией могут быть предоставлены в качестве информации, касающейся каждого из объектных сигналов. В этом случае после того, как задаются информация, указывающая объектный сигнал с наивысшей энергией, и абсолютное значение уровня энергии объектного сигнала с наивысшей энергией, могут быть легко определены уровни энергии других объектных сигналов не с наивысшей энергией.

Уровень энергии объектного сигнала с наивысшей энергией требуется для объединения множества потоков битов в один поток битов, выполняемого в узле управления многосторонней связью (MCU). Тем не менее, в большинстве случаев уровень энергии объектного сигнала с наивысшей энергией необязателен, поскольку абсолютное значение уровня энергии объектного сигнала с наивысшей энергией может быть легко определено из отношений уровней энергии других объектных сигналов не с наивысшей энергией к уровню энергии объектного сигнала с наивысшей энергией.

Например, предположим, что имеется четыре объектных сигнала А, В, С и D, принадлежащих предварительно определенной параметрической полосе, и что объектный сигнал А является объектным сигналом с наивысшей энергией. Далее, энергия Е предварительно определенной параметрической полосы и абсолютное значение Е уровня энергии объектного сигнала А удовлетворяют уравнению (2):

**Уравнение 2**

$$E_p = E_A + (a + b + c)E_A$$

$$E_A = \frac{E_p}{1 + a + b + c}$$

При этом  $a$ ,  $b$  и  $c$  соответственно указывают отношения уровня энергии объектных сигналов  $B$ ,  $C$  и  $D$  к уровню энергии объектного сигнала. Ссылаясь на уравнение (2), можно вычислить абсолютное значение  $E_A$  уровня энергии объектного сигнала  $A$  на основе отношений  $a$ ,  $b$  и  $c$  и энергии  $E_p$  предварительно определенной параметрической полосы. Следовательно, если только нет необходимости объединить множество битовых потоков в один битовый поток с помощью MCU, абсолютное значение  $E_A$  уровня энергии объектного сигнала  $A$ , возможно, не потребуется включать в поток битов. Информация, указывающая то, включено ли абсолютное значение  $E_A$  уровня энергии объектного сигнала  $A$  в поток битов, может быть включена в заголовок потока битов, тем самым, уменьшая размер потока битов.

С другой стороны, если есть необходимость объединить множество потоков битов в один поток битов с помощью MCU, уровень энергии объектного сигнала с наивысшей энергией является обязательным. В этом случае сумма уровней энергии, вычисленных на основе отношений уровней энергии объектных сигналов не с наивысшей энергией к уровню энергии объектного сигнала с наивысшей энергией, может не быть такой же, как уровень энергии сигнала понижающего микширования, полученного посредством понижающего микширования всех объектных сигналов. Например, когда уровень энергии сигнала понижающего микширования равен 100, сумма вычисленных уровней энергии может составлять 98 или 103 вследствие, к примеру, ошибок, возникающих в ходе операций квантования и деквантования. Чтобы разрешить это, разность между уровнем энергии сигнала понижающего микширования и суммой вычисленных уровней энергии может быть надлежащим образом компенсирована посредством умножения каждого из вычисленных уровней энергии на предварительно определенный коэффициент. Если уровень энергии сигнала понижающего микширования составляет  $X$ , а сумма вычисленных уровней энергии составляет  $Y$ , каждый из вычисленных уровней энергии может быть умножен на  $X/Y$ . Если разность между уровнем энергии сигнала понижающего микширования и суммой вычисленных уровней энергии не компенсирована, эти ошибки квантования могут быть включены в параметрические полосы и кадры, тем самым вызывая искажения сигнала.

Следовательно, информация, указывающая то, какой из множества объектных сигналов имеет наибольшее абсолютное значение энергии в предварительно определенной параметрической полосе, является обязательной. Эта информация может быть представлена посредством числа битов. Число битов, требуемых для указания того, какой из множества объектных сигналов имеет наибольшее абсолютное значение энергии в предварительно определенной параметрической полосе, варьируется согласно числу объектных сигналов. По мере того как число объектных сигналов возрастает, число битов, требуемых для указания того, какой из множества объектных сигналов имеет наибольшее абсолютное значение энергии в предварительно определенной параметрической полосе, возрастает. С другой стороны, по мере того как число объектных сигналов снижается, число битов, требуемых для указания того, какой из множества объектных сигналов имеет наибольшее абсолютное значение энергии в предварительно определенной параметрической полосе, уменьшается. Предварительно определенное число битов может быть назначено заранее для указания того, какой из множества объектных сигналов имеет наибольшее абсолютное значение энергии в предварительно определенной параметрической полосе. Альтернативно, число битов для указания того, какой из множества объектных сигналов имеет наибольшее абсолютное

значение энергии в предварительно определенной параметрической полосе, может быть определено на основе определенной информации.

Размер информации, указывающей то, какой из множества объектных сигналов имеет наибольшее абсолютное значение энергии в каждой параметрической полосе, может быть уменьшен посредством использования того же способа, что и использован для того, чтобы уменьшать размер информации CLD, ICC и CPC для использования в OTT- и/или TTT-модулях многоканального кодека, например, посредством использования временного дифференциального метода, частотного дифференциального метода или метода контрольного кодирования.

Чтобы указать то, какой из множества объектных сигналов имеет наибольшее абсолютное значение энергии в каждой параметрической полосе, может быть использована оптимизированная таблица Хаффмана. В этом случае информация, указывающая то, в каком порядке уровни энергии объектных сигналов сравниваются с уровнем энергии какого бы то ни было объектного сигнала, имеющего наибольшую абсолютную энергию, может потребоваться. Например, если имеется пять объектных сигналов (т.е. с первого по пятый объектные сигналы) и третий объектный сигнал является объектным сигналом с наивысшей энергией, может быть предоставлена информация, касающаяся третьего объектного сигнала. Далее отношения уровней энергии первого, второго, четвертого и пятого объектных сигналов к уровню энергии третьего объектного сигнала могут быть предоставлены различными способами, и это далее подробнее описывается.

Отношения уровней энергии первого, второго, четвертого и пятого объектных сигналов к уровню энергии третьего объектного сигнала могут предоставляться последовательно. Альтернативно, отношения уровней энергии четвертого, пятого, первого и второго объектных сигналов к уровню энергии третьего объектного сигнала могут предоставляться последовательно круговым способом. Далее информация, указывающая порядок, в котором предоставляются отношения уровней энергии первого, второго, четвертого и пятого объектных сигналов к уровню энергии третьего объектного сигнала, может быть включена в заголовок файла или может быть передана с интервалами числа кадров. Многоканальный кодек может определять информацию CLD и ICC на основе порядковых номеров OTT-модулей. Аналогично, необходима информация, указывающая то, как каждый объект отображается на поток битов.

В случае многоканального кодека информация, касающаяся сигналов, соответствующих каждому каналу, может быть идентифицирована посредством порядковых номеров OTT- или TTT-модулей. Согласно способу кодирования основывающегося на объектах аудио, если имеется N объектных сигналов, эти N объектных сигналов, возможно, потребуется соответствующим образом пронумеровать. Тем не менее, иногда пользователю необходимо контролировать N объектных сигналов с помощью объектного декодера. В этом случае пользователю могут быть нужны не только порядковые номера N объектных сигналов, но также описания N объектных сигналов, такие как описания, указывающие то, что первый объектный сигнал соответствует голосу женщины и что второй объектный сигнал соответствует звуку пианино. Описания N объектных сигналов могут быть включены в поток битов как метаданные и затем переданы вместе с потоком битов. Более конкретно, описания N объектных сигналов могут быть предоставлены как текст или могут быть предоставлены с помощью кодовой таблицы или кодовых слов.

Иногда необходима информация корреляции, касающаяся корреляций между

объектными сигналами. Для этого могут быть вычислены корреляции между объектным сигналом с наивысшей энергией и другими объектными сигналами не с наивысшей энергией. В этом случае одно корреляционное значение может быть  
5 указано для всех объектных сигналов, что сравнимо с использованием одного значения ИСС во всех ОТТ-модулях.

Если объектные сигналы являются стереосигналами, необходима информация отношений энергии левого канала к энергии правого канала для объектных сигналов и ИСС. Отношения энергии левого канала к энергии правого канала для объектных  
10 сигналов могут быть вычислены с помощью того же способа, что и использован для того, чтобы вычислять уровни энергии множества объектных сигналов на основе абсолютного значения уровня энергии того из объектных сигналов, который является объектным сигналом с наивысшей энергией, и отношений уровней энергии других  
15 объектных сигналов не с наивысшей энергией к уровню энергии объектного сигнала с наивысшей энергией. Например, если абсолютные значения уровней энергии левого и правого каналов объектного сигнала с наивысшей энергией равны А и В соответственно и отношение уровня энергии левого канала объектного сигнала не с наивысшей энергией к А и отношение уровня энергии правого канала объектного  
20 сигнала не с наивысшей энергией к В равны х и у соответственно, уровни энергии левого и правого каналов объектного сигнала не с наивысшей энергией могут быть вычислены как  $A \cdot x$  и  $B \cdot y$ . Таким образом, может быть вычислено отношение энергии левого канала к энергии правого канала объектного стереосигнала.

Абсолютное значение уровня энергии объектного сигнала с наивысшей энергией и  
25 отношения уровней энергии других объектных сигналов не с наивысшей энергией к уровню энергии объектного сигнала с наивысшей энергией также может быть использовано, когда объектные сигналы являются моносигналами, сигнал понижающего микширования, полученный посредством объектных моносигналов,  
30 является стереосигналом, и объектные моносигналы включены в оба канала стереосигнала понижающего микширования. В этом случае соотношение энергии части каждого объектного моносигнала, включенного в левый канал стереосигнала понижающего микширования, и энергии части каждого объектного моносигнала, включенного в правый канал стереосигнала понижающего микширования, и  
35 корреляционная информация необходимы, и это непосредственно применяется к объектным стереосигналам. Если объектный моносигнал включен в L- и R-каналы стереосигнала понижающего микширования, компоненты L- и R-каналов объектного моносигнала могут иметь только разность каналов, и объектный моносигнал может  
40 иметь значение корреляции 1 по всем параметрическим полосам. В этом случае для того, чтобы уменьшить объем данных, информация, указывающая то, что объектный моносигнал имеет значение корреляции, равное 1, по всем параметрическим полосам, может быть дополнительно предоставлена. Далее, нет необходимости указывать значение корреляции, равное 1, для каждой из параметрических полос. Вместо этого  
45 значение корреляции, равное 1, может быть указано для всех параметрических полос.

В ходе формирования сигнала понижающего микширования посредством суммирования множества объектных сигналов может возникать отсечение. Чтобы разрешить это, сигнал понижающего микширования может быть умножен на  
50 предварительно заданное усиление, с тем, чтобы максимальный уровень сигнала понижающего микширования мог превышать порог отсечения. Предварительно заданное усиление может варьироваться во времени. Следовательно, необходима информация, касающаяся предварительно заданного усиления. Если сигнал

понижающего микширования является стереосигналом, различные значения усиления могут быть предоставлены для L- и R-каналов сигнала понижающего микширования, чтобы не допустить отсечения. Чтобы уменьшить объем передачи данных, различные значения усиления могут не передаваться отдельно. Вместо этого сумма различных значений усиления и соотношение различных значений усиления может передаваться. Далее можно снижать динамический диапазон и уменьшать объем передачи данных в сравнении со случаем передачи различных значений усиления отдельно.

Чтобы дополнительно уменьшить объем передачи данных, может быть предоставлен бит, указывающий то, произошло ли отсечение в ходе отделения сигнала понижающего микширования, в ходе суммирования множества объектных сигналов. Затем, только если определено, что отсечение произошло, могут быть переданы значения усиления. Эта информация отсечения может быть необходимой для предотвращения отсечения в ходе суммирования множества сигналов понижающего микширования, чтобы объединить множество потоков битов. Чтобы предотвратить отсечение, сумма множества сигналов понижающего микширования может быть умножена на обратное число предварительно заданного значения усиления для предотвращения отсечения.

Фиг.13-16 иллюстрируют схемы для пояснения различных способов конфигурирования основывающейся на объектах дополнительной информации. Варианты осуществления по фиг.13-16 могут быть применены не только к объектным моно- и стереосигналам, но также для многоканальных объектных сигналов.

Ссылаясь на фиг.13, многоканальный объектный сигнал (от ОБЪЕСТ А(СН1) по ОБЪЕСТ А(СНn)) вводится в объектный кодер 221. Затем объектный кодер 221 формирует сигнал понижающего микширования и дополнительную информацию на основе многоканального объектного сигнала (от ОБЪЕСТ А(СН1) до ОБЪЕСТ А(СНn)). Объектный кодер 223 принимает множество объектных сигналов от ОБЪЕСТ1 до ОБЪЕСТn и сигнал понижающего микширования, сформированный посредством объектного кодера 221, и формирует еще один сигнал понижающего микширования и еще одну дополнительную информацию на основе объектных сигналов от ОБJ.1 до ОБJ.N и принятого сигнала понижающего микширования. Мультиплексор 225 объединяет дополнительную информацию, формируемую посредством объектного кодера 221, и дополнительную информацию, формируемую посредством объектного кодера 223.

Ссылаясь на фиг.14, объектный кодер 233 формирует первый поток битов на основе многоканального объектного сигнала (от ОБЪЕСТ А(СН1) до ОБЪЕСТ А(СНn)). Далее объектный кодер 231 формирует второй поток битов на основе множества немногочанальных объектных сигналов от ОБЪЕСТ1 до ОБЪЕСТn. Затем объектный кодер 235 комбинирует первый и второй потоки битов в один поток битов посредством использования практически того же способа, что и использован для того, чтобы объединить множество потоков битов в один поток битов с помощью MCU.

Ссылаясь на фиг.15, многоканальный кодер 241 формирует сигнал понижающего микширования и основывающуюся на каналах дополнительную информацию на основе многоканального объектного сигнала (от ОБЪЕСТ А(СН1) до ОБЪЕСТ А(СНn)). Объектный кодер 243 принимает сигнал понижающего микширования, сформированный посредством многоканального кодера 241, и множество немногочанальных объектных сигналов от ОБЪЕСТ1 до ОБЪЕСТn и формирует объектный поток битов и дополнительную информацию на основе принятого сигнала понижающего микширования и объектных сигналов от ОБЪЕСТ1 до ОБЪЕСТn.

Мультиплексор 245 комбинирует основывающуюся на каналах дополнительную информацию, сформированную посредством многоканального кодера 241, и дополнительную информацию, сформированную посредством объектного кодера 243, и выводит результат комбинирования.

5 Ссылаясь на фиг.16, многоканальный кодер 253 формирует сигнал понижающего микширования и основывающуюся на каналах дополнительную информацию на основе многоканального объектного сигнала (от ОБЪЕСТ А(СН1) до ОБЪЕСТ А(СНn)). Объектный кодер 251 формирует сигнал понижающего микширования и  
10 дополнительную информацию на основе множества немногочанальных объектных сигналов от ОБЪЕСТ1 до ОБЪЕСТn. Объектный кодер 255 принимает сигнал понижающего микширования, сформированный посредством многоканального кодера 253, и сигнал понижающего микширования, формируемый посредством объектного кодера 251, и комбинирует принятые сигналы понижающего  
15 микширования. Мультиплексор 257 комбинирует дополнительную информацию, сформированную посредством объектного кодера 251, и основывающуюся на каналах дополнительную информацию, сформированную посредством многоканального кодера 243, и выводит результат комбинирования.

20 В случае кодирования основывающегося на объектах аудио в телеконференциях иногда необходимо объединить множество объектных потоков битов в один поток битов. Ниже подробно описывается объединение множества объектных потоков битов в один поток битов.

Фиг.17 иллюстрирует схему для пояснения объединения двух объектных потоков битов. Ссылаясь на фиг.17, когда два объектных потока битов объединяются в один объектный поток битов, дополнительная информация, такая как информация CLD и ИСС, представленная в двух объектных потоках битов, соответственно, может быть модифицирована. Два объектных потока битов могут быть объединены в один  
30 объектный поток битов просто посредством использования дополнительного ОТТ-модуля, т.е. одиннадцатого ОТТ-модуля, и использования дополнительной информации, такой как информация CLD и ИСС, предоставляемой посредством одиннадцатого ОТТ-модуля.

35 Древовидная конфигурационная информация каждого из двух объектных потоков битов должна быть объединена в интегрированную древовидную конфигурационную информацию, чтобы объединить два объектных потока битов в один объектный поток битов. Для этого, дополнительная конфигурационная информация, если имеется, сформированная посредством объединения двух объектных потоков битов, может быть модифицирована, индексы числа ОТТ-модулей, используемых для того,  
40 чтобы формировать два объектных потока битов, могут быть модифицированы, и только несколько дополнительных процессов, таких как процесс вычислений, выполняемый посредством одиннадцатого ОТТ-модуля, и понижающее микширование двух сигналов понижающего микширования в двух объектных  
45 потоках битов могут быть выполнены. Таким образом, два объектных потока битов могут быть легко объединены в один объектный поток битов без необходимости модифицировать информацию, касающуюся каждого из множества объектных сигналов, из которых исходят два объектных сигнала.

50 Ссылаясь на фиг.17, одиннадцатый ОТТ-модуль может быть необязательным. В этом случае два сигнала понижающего микширования двух объектных потоков битов могут быть использованы, как есть, в качестве двухканального сигнала понижающего микширования. Таким образом, два объектных потока битов могут быть объединены

в один поток битов без необходимости дополнительных вычислений.

Фиг.18 иллюстрирует схему для пояснения объединения двух или более независимых объектных потоков битов в один объектный поток битов, имеющий стереосигнал понижающего микширования. Ссылаясь на фиг.18, если два или более независимых потока битов имеют различное число параметрических полос, преобразование параметрических полос может выполняться для объектных потоков битов, так чтобы число параметрических полос одного из объектных потоков битов, имеющего меньше параметрических полос, могло быть увеличено, чтобы быть таким же, как число параметрических полос другого объектного потока битов.

Более конкретно, преобразование параметрической полосы может выполняться с помощью предварительно определенной таблицы преобразования. В этом случае преобразование параметрической полосы может выполняться с помощью простой линейной формулы.

Если имеются перекрывающиеся параметрические полосы, значения параметров могут надлежащим образом смешиваться с учетом величины, на которую перекрывающиеся параметрические полосы перекрывают друг друга. В ситуациях, когда низкая сложность приоритетна, преобразование параметрических полос может выполняться для двух объектных потоков битов, с тем, чтобы число параметрических полос одного из двух объектных потоков битов, имеющего больше параметрических полос, могло быть уменьшено так, чтобы быть таким же, как число параметрических полос другого объектного потока битов.

В вариантах осуществления по фиг.17 и 18 два или более независимых объектных потока битов могут быть объединены в интегрированный объектный поток битов без необходимости вычисления существующих параметров независимых объектных потоков битов. Тем не менее, в случае объединения множества сигналов понижающего микширования параметры, касающиеся сигналов понижающего микширования, возможно, потребуются вычислить снова посредством QMF/гибридного анализа.

Тем не менее, это вычисление требует значительного объема вычислений, тем самым подвергая риску преимущества вариантов осуществления по фиг.17 и 18. Следовательно, необходимо предложить способы извлечения параметров без необходимости QMF/гибридного анализа или синтеза, даже когда сигналы понижающего микширования микшируются с понижением. Для этого информация энергии, касающаяся энергий каждой параметрической полосы каждого сигнала понижающего микширования, может быть включена в объектный поток битов. Затем, когда сигналы понижающего микширования смикшированы с понижением, информация, такая как информация CLD, может быть легко вычислена на основе такой информации энергии без необходимости QMF/гибридного анализа или синтеза. Эта информация энергии может представлять наивысший уровень энергии для каждой параметрической полосы или абсолютное значение уровня энергии объектного сигнала с наивысшей энергией для каждой параметрической полосы. Объем вычислений дополнительно может быть снижен посредством использования значений ИСС, полученных из временной области для каждой параметрической полосы.

В ходе понижающего микширования множества сигналов понижающего микширования может возникать отсечение. Чтобы разрешить это, уровни сигналов понижающего микширования могут быть уменьшены. Если уровни сигналов понижающего микширования уменьшены, информация уровней, касающаяся уменьшенных уровней сигналов понижающего микширования, возможно, должна

быть включена в объектный поток битов. Информация уровня для недопущения отсечения может быть применена к каждому кадру объектного потока битов или может быть применена только к некоторым кадрам, в которых возникает отсечение. Уровни исходных сигналов понижающего микширования могут быть вычислены  
5 посредством обратного применения информации уровня для недопущения отсечения в ходе операции декодирования. Информация уровня для недопущения отсечения может быть вычислена во временной области и тем самым не должна быть подвергнута QMF/гибридному синтезу или анализу. Объединение множества объектных сигналов в  
10 один объектный поток битов может выполняться с помощью структуры, проиллюстрированной на фиг.12, и это далее подробнее описывается со ссылкой на фиг.19.

Фиг.19 иллюстрирует схему для пояснения объединения двух независимых объектных потоков битов в один объектный поток битов. Ссылаясь на фиг.19, первый  
15 модуль 261 формирует первый объектный поток битов, а второй модуль 263 формирует второй объектный поток битов. Далее, третий модуль 265 формирует третий поток битов посредством комбинирования первого и второго потока битов. В этом случае, если первый и второй объектные потоки битов включают в себя  
20 информацию по абсолютному значению уровня энергии объектного сигнала с наивысшей энергией для каждой параметрической полосы и отношениям уровней энергии других объектных сигналов не с наивысшим уровнем энергии к уровню энергии объектного сигнала с наивысшим уровнем энергии и информацию усиления,  
25 касающуюся значений усиления, которые умножены на сигналы понижающего микширования посредством первого и второго модулей 261 и 263, третий модуль 265 может формировать третий объектный поток битов просто посредством объединения первого и второго потоков битов без необходимости дополнительного вычисления или извлечения параметров.

Третий модуль 265 принимает множество сигналов понижающего микширования DOWNMIX\_A и DOWNMIX\_B. Третий модуль 265 преобразует  
30 сигналы понижающего микширования DOWNMIX\_A и DOWNMIX\_B в PCM-сигналы и суммирует PCM-сигналы, тем самым формируя один сигнал понижающего микширования. Тем не менее, в ходе этого процесса может возникать отсечение.  
35 Чтобы разрешить это, сигналы понижающего микширования DOWNMIX\_A и DOWNMIX\_B могут быть умножены на предварительно заданное значение усиления. Информация, касающаяся предварительно заданного значения усиления, может быть включена в третий объектный поток битов и передана вместе с третьим  
40 объектным потоком битов.

Ниже подробно описывается объединение множества объектных потоков битов в один поток битов. Ссылаясь на фиг.19, paramA может включать в себя информацию, касающуюся того, какой из множества объектных сигналов от ОБЪЕКТ1 до ОБЪЕКТn  
45 является объектным сигналом с наивысшей энергией, и отношений уровней энергии других объектных сигналов не с наивысшей энергией к уровню энергии объектного сигнала с наивысшей энергией. Аналогично, SIDE\_INFO может включать в себя SIDE\_INFO\_A, которое может включать в себя информацию, касающуюся того, какой из множества объектных сигналов от ОБЪЕКТ1 до ОБЪЕКТn является  
50 объектным сигналом с наивысшей энергией, и отношений уровней энергии других объектных сигналов не с наивысшей энергией к уровню энергии объектного сигнала с наивысшей энергией.

SIDE\_INFO\_A и SIDE\_INFO\_B могут быть включены параллельно в один поток

битов, как проиллюстрировано на фиг.20. В этом случае дополнительно может быть предоставлен бит, указывающий то, существует более одного потока битов параллельно.

5 Ссылаясь на фиг.20, чтобы указать то, является ли предварительно определенный поток битов интегрированным потоком битов, включающим в себя более одного потока битов, информация, указывающая то, является ли предварительно  
определенный поток битов интегрированным потоком битов, информация,  
10 касающаяся числа потоков битов, если есть, включенных в предварительно  
определенный поток битов, и информация, касающаяся исходных позиций потоков битов, если есть, включенных в предварительно определенный поток битов, может  
быть предоставлена в заголовке предварительно определенного потока битов, за  
15 которым следует более одного потока битов, если есть, в предварительно  
определенный поток битов. В этом случае декодер может определять то, является ли  
предварительно определенный поток битов интегрированным потоком битов,  
включающим в себя более одного потока битов, посредством анализа информации в  
заголовке предварительно определенного потока битов. Этот тип объединения  
20 потоков битов не требует дополнительных процессов, помимо добавления нескольких  
идентификаторов в поток битов. Тем не менее, такие идентификаторы должны  
предоставляться с интервалами в число кадров. Помимо этого, этот тип способа  
объединения потока битов требует от декодера определять то, является ли каждый  
поток битов, который принимает декодер, интегрированным потоком битов.

25 В качестве альтернативы вышеописанному способу объединения потоков битов,  
множество потоков битов мог быть объединены в один поток битов таким образом,  
чтобы декодер не могут распознать, является или нет один поток битов  
интегрированным потоком битов. Это подробно описывается ниже со ссылкой на  
фиг.21.

30 Ссылаясь на фиг.21, сравниваются уровень энергии объектного сигнала с  
наивысшей энергией, представленного посредством SIDE\_INFO\_A, и уровень энергии  
объектного сигнала с наивысшей энергией, представленного  
35 посредством SIDE\_INFO\_B. Затем тот из двух объектных сигналов, который имеет  
наивысший уровень энергии, определяется как объектный сигнал с наивысшей  
энергией интегрированного потока битов. Например, если уровень энергии  
объектного сигнала с наивысшей энергией, представленного  
40 посредством SIDE\_INFO\_A, выше уровня энергии объектного сигнала с наивысшей  
энергией, представленного посредством SIDE\_INFO\_B, объектный сигнал с наивысшей  
энергией, представленный посредством SIDE\_INFO\_A, может становиться объектным  
сигналом с наивысшей энергией интегрированного потока битов. Затем информация  
энергетического соотношения SIDE\_INFO\_A может быть использована в  
интегрированном потоке битов как есть, тогда как информация энергетического  
45 соотношения SIDE\_INFO\_B может быть умножена на отношение уровней энергии.

Далее, информация энергетического соотношения того из SIDE\_INFO\_A  
и SIDE\_INFO\_B, который включает в себя информацию, касающуюся объектного  
сигнала с наивысшей энергией интегрированного потока битов, может быть  
использована в качестве интегрированного потока битов, и информацию  
50 энергетического соотношения объектного сигнала с наивысшей энергией,  
представленного посредством Param A, и объектного сигнала с наивысшей энергией,  
представленного посредством SIDE\_INFO\_B. Этот способ влечет за собой пересчет  
информации энергетического соотношения SIDE\_INFO\_B. Тем не менее, пересчет

информации энергетического соотношения SIDE\_INFO\_V является относительно несложным. В этом способе декодер может не иметь возможности определять то, является поток битов, который он принимает, интегрированным потоком битов, включающим в себя более одного потока битов, и тем самым может быть

5

использован типичный способ декодирования. Два объектных потока битов, включающих в себя стереосигналы понижающего микширования, легко могут быть объединены в один объектный поток битов без необходимости пересчета информации, касающейся объектных сигналов, посредством

10

15

использования практически такого же способа, что использован для того, чтобы объединять потоки битов, включающие в себя моносигналы понижающего микширования. В объектном потоке битов информация, касающаяся древовидной структуры, которая низводит объектный сигнал, следует после информации объектных сигналов, получаемой из каждой ветви (т.е. каждого модуля) древовидной

20

структуры. Объектные потоки битов описаны выше с учетом того, что определенный объект распределяется не только левому каналу или правому каналу стереосигнала понижающего микширования. Тем не менее, объектные сигналы, в общем, распределяются между обоими каналами стереосигнала понижающего микширования. Следовательно, далее подробно описывается то, как формировать объектный поток битов на основе потоков битов, которые распределяются между двумя каналами стереосигнала понижающего микширования.

25

30

### Уравнение 3

35

$$Eng_{ObjL} = \frac{a}{a+b} Eng_{Obj}$$

$$Eng_{ObjR} = \frac{b}{a+b} Eng_{Obj}$$

40

Если объектный сигнал распределяется между L- и R-каналами стереосигнала понижающего микширования, может дополнительно потребоваться информация соотношения распределения по каналам, касающаяся соотношения (a:b), с которым объектный сигнал распределяется между L- и R-каналами. Далее информация, касающаяся объектного сигнала, такая как информация CLD и ICC, может быть

45

вычислена посредством выполнения понижающего микширования с помощью ОТТ-модулей для L- и R-каналов стереосигнала понижающего микширования, и это подробнее описывается далее со ссылкой на фиг.23.

50

Ссылаясь на фиг.23, после того как информация CLD и ICC, полученная из множества ОТТ-модулей в ходе операции понижающего микширования, и информация соотношения распределения между каналами каждого из множества объектных сигналов предоставлена, можно вычислять многоканальный поток битов, который

адаптивно варьируется согласно любой модификации, выполненной в информации позиции объектов и информации конфигурации воспроизведения посредством конечного пользователя. Помимо того, если стереосигнал понижающего микширования должен быть обработан посредством предварительной обработки  
5 понижающего микширования, можно получить информацию, касающуюся того, как стереосигнал понижающего микширования обработан посредством предварительной обработки понижающего микширования, и передавать полученную информацию в препроцессор. Т.е. если информация соотношения распределения по каналам каждого  
10 из множества объектных сигналов не предоставлена, нет способа вычислять многоканальный поток битов и получать информацию, требуемую для работы препроцессора. Информация соотношения распределения между каналами объектного сигнала может быть представлена как соотношение двух целых чисел или скалярная величина (в дБ).

15 Как описано выше, если объектный сигнал распределен между двумя каналами стереосигнала понижающего микширования, информация соотношения распределения между каналами объектного сигнала может требоваться. Информация соотношения распределения между каналами может иметь фиксированное значение, указывающее  
20 соотношение, с которым объектный сигнал распределен между двумя каналами стереосигнала понижающего микширования. Альтернативно, информация соотношения распределения между каналами объектного сигнала может варьироваться от одной полосы частот до другой полосы частот объектного сигнала, особенно когда информация соотношения распределения между каналами  
25 используется как информация ИСС. Если стереосигнал понижающего микширования получается посредством усложненной операции понижающего микширования, т.е. если объектный сигнал принадлежит двум каналам стереосигнала понижающего микширования и микшируется с понижением посредством варьирования  
30 информации ИСС от одной полосы частот до другой полосы частот объектного сигнала, подробное описание понижающего микширования объектного сигнала может дополнительно потребоваться, чтобы декодировать окончательно подготовленный посредством рендеринга объектный сигнал. Этот вариант осуществления может быть применен ко всем возможным объектным структурам,  
35 которые уже описаны.

Далее подробно описывается предварительная обработка со ссылкой на фиг.24-27. Если сигнал понижающего микширования, введенный в объектный декодер, является стереосигналом, входной сигнал понижающего микширования, возможно,  
40 потребуется предварительно обработать перед вводом в многоканальный декодер объектного декодера, поскольку многоканальный декодер не может преобразовывать сигнал, принадлежащий левому каналу входного сигнала понижающего микширования, в правый канал. Следовательно, чтобы конечный пользователь сдвинул позицию объектного сигнала, принадлежащего левому каналу входного  
45 сигнала понижающего микширования, к правому каналу, входной сигнал понижающего микширования, возможно, должен быть предварительно обработан, и предварительно обработанный сигнал понижающего микширования может быть введен в многоканальный декодер.

50 Предварительная обработка стереосигнала понижающего микширования может быть выполнена посредством получения информации предварительной обработки из объектного потока битов и из матрицы рендеринга и надлежащей обработки стереосигнала понижающего микширования согласно информации обработки, и это

далее подробно описывается.

Фиг.24 иллюстрирует схему для пояснения того, как конфигурировать стереосигнал понижающего микширования на основе четырех объектных сигналов от ОБЪЕСТ1 до ОБЪЕСТ4. Ссылаясь на фиг.24, первый объектный сигнал ОБЪЕСТ1 распределяется между L- и R-каналами с соотношением a:b, второй объектный сигнал ОБЪЕСТ2 распределяется между L- и R-каналами с соотношением c:d, третий объектный сигнал ОБЪЕСТ3 распределяется только L-каналу, и четвертый объектный сигнал ОБЪЕСТ4 распределяется только R-каналу. Информация, такая как CLD и ICC, может быть сформирована посредством передачи каждого из с первого по четвертый объектного сигнала от ОБЪЕСТ1 по ОБЪЕСТ4 посредством ряда ОТТ, и сигнал понижающего микширования может быть сформирован на основе сформированной информации.

Предположим, что конечный пользователь получает матрицу рендеринга посредством надлежащего задания позиций и уровней с первого по четвертый объектных сигналов от ОБЪЕСТ1 до ОБЪЕСТ4 и что имеется пять каналов. Матрица рендеринга может быть представлена посредством уравнения (4):

**Уравнение 4**

$$\begin{bmatrix} 30 & 10 & 20 & 30 & 10 \\ 10 & 30 & 20 & 10 & 30 \\ 22 & 22 & 22 & 22 & 22 \\ 21 & 21 & 31 & 11 & 11 \end{bmatrix}$$

Ссылаясь на уравнение (4), когда сумма из пяти коэффициентов в каждой из четырех строк равна предварительно заданному опорному значению, т.е. 100, определяется то, что уровень соответствующего объектного сигнала не варьировался. Величина, на которую сумма пяти коэффициентов в каждой из четырех строк отличается от предварительно заданного опорного значения, может быть величиной (в дБ), на которую варьировался уровень соответствующего объектного сигнала. Первый, второй, третий, четвертый и пятый столбцы в матрице рендеринга уравнения (4) представляют FL-, FR-, C-, RL- и RR-каналы соответственно.

Первая строка матрицы рендеринга по уравнению (4) соответствует первому объектному сигналу ОБЪЕСТ1 и имеет всего пять коэффициентов, т.е. 30, 10, 20, 30 и 10. Поскольку сумма пяти коэффициентов первой строки составляет 100, определяется, что уровень первого объектного сигнала ОБЪЕСТ1 не варьировался, и что только пространственная позиция первого объектного сигнала ОБЪЕСТ1 изменилась. Даже если пять коэффициентов первой строки представляют различные каналные направления, они могут крупно классифицироваться на два канала: L- и R-каналы. В таком случае соотношение, с которым первый объектный сигнал ОБЪЕСТ1 распределяется между L- и R-каналами, может быть вычислено как  $70\%=(30+30+20)*0,5$ : $30\%=(10+10+20)*0,5$ . Следовательно, матрица рендеринга по уравнению (4) указывает то, что уровень первого объектного сигнала ОБЪЕСТ1 не варьировался, и что первый объектный сигнал ОБЪЕСТ1 распределен между L- и R-каналами с соотношением 70%:30%. Если сумма пяти коэффициентов любой из строк матрицы рендеринга по уравнению (4) меньше или не больше 100, может быть определено, что уровень соответствующего объектного сигнала изменился, и затем соответствующий объектный сигнал может быть обработан посредством предварительной обработки или может быть преобразован и передан как ADG.

Чтобы предварительно обрабатывать сигналы понижающего микширования,

соотношение, с которым сигналы понижающего микширования распределяются между параметрическими полосами, из которых параметры извлекаются из сигналов, полученных посредством выполнения QMF/гибридного преобразования сигналов понижающего микширования, может быть вычислено, и сигналы понижающего микширования могут быть перераспределены между параметрическими полосами согласно настройке матрицы рендеринга. Далее подробно описываются различные способы перераспределения сигналов понижающего микширования между параметрическими полосами.

В первом способе перераспределения сигналы понижающего микширования L- и R-канала декодируются отдельно с помощью их соответствующей дополнительной информации (такой как информация CLD и ICC) и с помощью практически того же способа, что и использован посредством многоканального кодека. Далее объектные сигналы, распределенные между сигналами понижающего микширования L- и R-канала, восстанавливаются. Чтобы снизить объем вычислений, сигналы понижающего микширования L- и R-канала могут быть декодированы только с помощью информации CLD. Соотношение, с которым каждый из восстановленных объектных сигналов распределяется между сигналами понижающего микширования L- и R-канала, может быть определено на основе дополнительной информации.

Каждый из восстановленных объектных сигналов может быть перераспределен между сигналами понижающего микширования L- и R-канала согласно матрице рендеринга. Далее перераспределенные объектные сигналы микшируются с понижением на поканальной основе посредством ОТТ-модулей, тем самым, завершая предварительную обработку. Вкратце, первый способ перераспределения приспособливает тот же способ, что и используемый многоканальным кодеком. Тем не менее, первый способ перераспределения требует стольких процессов декодирования, сколько объектных сигналов для каждого канала, и требует процесса перераспределения и процесса основывающегося на каналах понижающего микширования.

Во втором способе перераспределения, в отличие от первого способа перераспределения, объектные сигналы не восстанавливаются из L- и R-сигналов понижающего микширования. Вместо этого, каждый из L- и R-сигналов понижающего микширования делится на две части: одна часть L\_L или R\_R, которая должна быть оставлена в соответствующем канале, и другая часть L\_R или R\_L, которая должна быть перераспределена, как проиллюстрировано на фиг.25. Ссылаясь на фиг.25, L\_L указывает часть сигнала понижающего микширования L-канала, которая должна быть левой в L-канале, а L\_R указывает часть сигнала понижающего микширования L-канала, которая должна быть добавлена к R-каналу. Аналогично, R\_R указывает часть сигнала понижающего микширования R-канала, которая должна быть оставлена в R-канале, а R\_L указывает часть сигнала понижающего микширования R-канала, которая должна быть добавлена к L-каналу. Каждый из сигналов понижающего микширования L- и R-канала может быть поделен на две части (L\_L и L\_R или R\_R и R\_L) согласно соотношению, с которым каждый объектный сигнал распределяется между L- и R-сигналами понижающего микширования, как задано посредством уравнения (2), и соотношению, с которым каждый объектный сигнал должен быть распределен между предварительно обработанными L и R в L- и R-каналах, как задано посредством уравнения (3). Следовательно, может быть определено, что сигналы понижающего микширования L- и R-каналов должны быть перераспределены между предварительно

обработанными L и R в L- и R-каналах посредством сравнения соотношения, с которым каждый объектный сигнал распределяется между L- и R-сигналами понижающего микширования, и соотношения, с которым каждый объектный сигнал должен быть распределен между предварительно обработанными L и R в L- и R-каналах.

Разделение сигнала L-канала на сигналы L\_L и L\_R согласно предварительно заданному энергетическому соотношению описано выше. После того как сигнал L-канала разделен на сигналы L\_L и L\_R, возможно, должно быть определено ICC между сигналами L\_L и L\_R. ICC между сигналами L\_L и L\_R могут быть легко определены на основе информации ICC, касающейся объектных сигналов. Т.е. ICC между сигналами L\_L и L\_R может быть определено на основе соотношения, с которым каждый объектный сигнал распределяется между сигналами L\_L и L\_R.

Далее подробно описывается второй способ перераспределения понижающего микширования. Предположим, что сигналы L и R понижающего микширования L- и R-каналов получены посредством способа, показанного на фиг.24, и что первый, второй, третий и четвертый объектные сигналы ОБЪЕКТ1, ОБЪЕКТ2, ОБЪЕКТ3 и ОБЪЕКТ4 распределяются между сигналами L и R понижающего микширования L- и R-каналов с соотношениями 1:2, 2:3, 1:0 и 0:1 соответственно. Множество объектных сигналов может быть смикшировано с понижением посредством ряда ОТТ-модулей, и информация, такая информация CLD и ICC, может быть получена из понижающего микширования объектных сигналов.

Пример матрицы рендеринга, установленной для с первого по четвертый объектных сигналов от ОБЪЕКТ1 по ОБЪЕКТ4, такой, как представлено посредством уравнения (4). Матрица рендеринга включает в себя информацию местоположения с первого по четвертый объектных сигналов от ОБЪЕКТ1 до ОБЪЕКТ4. Таким образом, предварительные сигналы L и R понижающего микширования L- и R-каналов могут быть получены посредством выполнения предварительной обработки с помощью матрицы рендеринга. То, как установить и интерпретировать матрицу рендеринга, уже описано выше со ссылкой на уравнение (3).

Соотношение, с которым каждый из с первого по четвертый объектного сигнала от ОБЪЕКТ1 до ОБЪЕКТ4 распределяется между предварительно обработанными сигналами L и R понижающего микширования L- и R-каналов, может быть вычислено так, как указано посредством уравнения (5):

#### Уравнение 5

Объект 1:

$$Eng_{Obj1_L} = 30 + 30 + 20 * 0.5 = 70, Eng_{Obj1_R} = 10 + 10 + 20 * 0.5 = 30$$

$$Eng_{Obj1_L} : Eng_{Obj1_R} = 70 : 30$$

Объект 2:

$$Eng_{Obj2_L} = 10 + 10 + 20 * 0.5 = 30, Eng_{Obj2_R} = 30 + 30 + 20 * 0.5 = 70$$

$$Eng_{Obj2_L} : Eng_{Obj2_R} = 30 : 70$$

Объект 3:

$$Eng_{Obj3_L} = 22 + 22 + 22 * 0.5 = 55, Eng_{Obj3_R} = 22 + 22 + 22 * 0.5 = 55$$

$$Eng_{Obj3_L} : Eng_{Obj3_R} = 55 : 55$$

Объект 4:

$$Eng_{Obj4_L} = 21 + 11 + 31 * 0.5 = 47.5, Eng_{Obj4_R} = 21 + 11 + 31 * 0.5 = 47.5$$

$$Eng_{Obj4_L} : Eng_{Obj4_R} = 47.5 : 47.5$$

Соотношение, с которым каждый из с первого по четвертый объектного сигнала от ОБЪЕСТ1 до ОБЪЕСТ4 распределяется между сигналами L и R понижающего микширования L- и R-каналов, может быть вычислено так, как указано посредством уравнения (6):

#### Уравнение 6

Объект 1:

$$Eng_{Obj1_L} : Eng_{Obj1_R} = 1 : 2$$

Объект 2:

$$Eng_{Obj2_L} : Eng_{Obj2_R} = 2 : 3$$

Объект 3:

$$Eng_{Obj3_L} : Eng_{Obj3_R} = 1 : 0$$

Объект 4:

$$Eng_{Obj4_L} : Eng_{Obj4_R} = 0 : 1$$

Ссылаясь на уравнение (5), сумма части третьего объектного сигнала ОБЪЕСТ3, распределяемого для предварительно обработанного сигнала L понижающего микширования L-канала, и части третьего объектного сигнала ОБЪЕСТ3, распределяемого для сигнала R понижающего микширования R-канала, равна 110, и таким образом, определяется то, что уровень третьего объектного сигнала ОБЪЕСТ3 увеличился на 10. С другой стороны, сумма части четвертого объектного сигнала ОБЪЕСТ4, распределяемого для предварительно обработанного сигнала L понижающего микширования L-канала, и части четвертого объектного сигнала ОБЪЕСТ4, распределяемого для сигнала R понижающего микширования R-канала, равна 85, и таким образом, определяется то, что уровень четвертого объектного сигнала ОБЪЕСТ4 снизился на 5. Если матрица рендеринга для с первого по четвертый объектных сигналов от ОБЪЕСТ1 до ОБЪЕСТ4 имеет опорное значение в 100, и величина, на которую сумма коэффициентов в каждой из строк матрицы рендеринга отличается от опорного значения в 100, представляет величину (единицы - дБ), на которую варьировался уровень соответствующего объектного сигнала, может быть определено то, что уровень третьего объектного сигнала ОБЪЕСТ3 увеличился на 10 дБ, и что уровень четвертого объектного сигнала ОБЪЕСТ4 уменьшился на 5 дБ.

Уравнения (5) и (6) могут быть перекомпонованы в уравнение (7):

#### Уравнение 7

Объект 1:

$$Eng_{Obj1_L} : Eng_{Obj1_R} = 33.3 : 66.7 \quad Eng_{Obj1_L} : Eng_{Obj1_R} = 70 : 30$$

Объект 2:

$$Eng_{Obj2_L} : Eng_{Obj2_R} = 40 : 60 \quad Eng_{Obj2_L} : Eng_{Obj2_R} = 30 : 70$$

Объект 3:

$$Eng_{Obj3_L} : Eng_{Obj3_R} = 100 : 0 \quad Eng_{Obj3_L} : Eng_{Obj3_R} = 50 : 50$$

Объект 4:

$$Eng_{Obj4_L} : Eng_{Obj4_R} = 0 : 100 \quad Eng_{Obj4_L} : Eng_{Obj4_R} = 50 : 50$$

Уравнение (7) сравнивает соотношение, с которым каждый из с первого по четвертый объектного сигнала от ОБЪЕСТ1 до ОБЪЕСТ4 распределяется между сигналами понижающего микширования L- и R-канала до предварительной обработки, и сравнивает соотношение, с которым каждый из с первого по четвертый

объектных сигналов от ОБЪЕСТ1 до ОБЪЕСТ4 распределяется между сигналами понижающего микширования L- и R-канала после предварительной обработки. Следовательно, посредством использования уравнения (7) можно легко определить то, какая часть каждого с первого по четвертый объектных сигналов от ОБЪЕСТ1 до ОБЪЕСТ4 должна быть перераспределена посредством предварительной обработки. Например, ссылаясь на (7), соотношение, с которым второй объектный сигнал ОБЪЕСТ2 распределяется между сигналами понижающего микширования L- и R-канала, изменяется с 40:60 до 30:70, и таким образом, может быть определено то, что одна четвертая (25%) части второго объектного сигнала ОБЪЕСТ2, ранее распределенная сигналу понижающего микширования L-канала, должна быть сдвинута к сигналу понижающего микширования R-канала. Это может стать более очевидным посредством ссылки на уравнение (8):

**Уравнение 8**

ОБЪЕСТ1: 55% части ОБЪЕСТ1, ранее распределенной R, должно быть сдвинуто к L  
 ОБЪЕСТ2: 25% части ОБЪЕСТ1, ранее распределенной L, должно быть сдвинуто к R  
 ОБЪЕСТ3: 50% части ОБЪЕСТ1, ранее распределенной L, должно быть сдвинуто к R  
 ОБЪЕСТ4: 50% части ОБЪЕСТ1, ранее распределенной R, должно быть сдвинуто к L.

Посредством использования уравнения (8) сигналы L\_L, L\_R, R\_L и R\_R по фиг.25 могут быть представлены так, как указано посредством уравнения (9):

**Уравнение 9**

$$Eng_{L\_L} = Eng_{Obj1_L} + 0.75 \cdot Eng_{Obj2_L} + 0.5 \cdot Eng_{Obj3}$$

$$Eng_{L\_R} = 0.25 \cdot Eng_{Obj2_L} + 0.5 \cdot Eng_{Obj3}$$

$$Eng_{R\_L} = 0.55 \cdot Eng_{Obj1_R} + 0.5 \cdot Eng_{Obj4}$$

$$Eng_{R\_R} = 0.45 \cdot Eng_{Obj1_R} + Eng_{Obj2_R} + 0.5 \cdot Eng_{Obj4}$$

Значение каждого объектного сигнала в уравнении (9) может быть представлено как соотношение, с которым соответствующий объектный сигнал распределяется между L- и R-каналами, посредством использования деквантованной информации CLD, предоставленной посредством ОТТ-модуля, как указано посредством уравнения (10):

**Уравнение 10**

$$Eng_{Obj1_L} = \frac{\frac{CLD2}{1+10^{-10}}}{\frac{CLD2}{1+10^{-10}}} \cdot \frac{\frac{CLD1}{1+10^{-10}}}{\frac{CLD1}{1+10^{-10}}} \cdot Eng_L, \quad Eng_{Obj2_L} = \frac{\frac{CLD2}{1+10^{-10}}}{\frac{CLD2}{1+10^{-10}}} \cdot \frac{1}{\frac{CLD1}{1+10^{-10}}} \cdot Eng_L$$

$$Eng_{Obj1_R} = \frac{\frac{CLD4}{1+10^{-10}}}{\frac{CLD4}{1+10^{-10}}} \cdot \frac{\frac{CLD3}{1+10^{-10}}}{\frac{CLD3}{1+10^{-10}}} \cdot Eng_R, \quad Eng_{Obj2_R} = \frac{\frac{CLD4}{1+10^{-10}}}{\frac{CLD4}{1+10^{-10}}} \cdot \frac{1}{\frac{CLD3}{1+10^{-10}}} \cdot Eng_R$$

$$Eng_{Obj3} = \sqrt{\frac{1}{1+10^{-10}}} \cdot Eng_L, \quad Eng_{Obj4} = \frac{1}{1+10^{-10}} \cdot Eng_R$$

Информация CLD, используемая в каждом блоке анализа по фиг.25, может быть определена так, как указано посредством уравнения (11):

**Уравнение 11**

$$CLD_{pars1} = 10 \log_{10} \left( \frac{L_L + \varepsilon}{L_R + \varepsilon} \right)$$

$$CLD_{pars2} = 10 \log_{10} \left( \frac{R_L + \varepsilon}{R_R + \varepsilon} \right)$$

где  $\varepsilon$  - константа для избежания деления на ноль, например, 96дБ ниже максимального ввода сигнала.

Таким образом, информация CLD и ICC, используемая в блоке анализа для формирования сигналов L\_L и L\_R на основе сигнала понижающего микширования L-канала, может быть определена, и информация CLD и ICC, используемая в блоке анализа для формирования сигналов R\_L и R\_R на основе сигнала понижающего микширования R-канала, также может быть определена. После того как сигналы L\_L, L\_R, R\_L и R\_R получены, как проиллюстрировано на фиг.25, сигналы L\_R и R\_R могут быть суммированы, тем самым получая предварительно обработанный стереосигнал понижающего микширования. Если конечный канал является стереоканалом, сигналы понижающего микширования L- и R-каналов, полученные посредством предварительной обработки, могут быть выведены. В этом случае должно быть отрегулировано варьирование, если имеет место, уровня каждого объектного сигнала. Для этого предварительно определенный модуль, который выполняет функции ADG-модуля, может быть дополнительно предоставлен. Информация для регулирования уровня каждого объектного сигнала может быть вычислена с помощью того же способа, что используемый для того, чтобы вычислять информацию ADG, и это подробнее описывается далее. Альтернативно, уровень каждого объектного сигнала может быть отрегулирован в ходе операции предварительной обработки. В этом случае регулирование уровня каждого объектного сигнала может выполняться с помощью того же способа, что и используемый для того, чтобы обрабатывать ADG. Альтернативно варианту осуществления по фиг.25, операция декорреляции может выполняться посредством декоррелятора и микшера, а не модулей анализа PARSING 1 и PARSING 2, как проиллюстрировано на фиг.26, чтобы регулировать корреляцию между сигналами L и R, полученными посредством микширования. Ссылаясь на фиг.26, Pre\_L и Pre\_R указывают сигналы L- и R-канала, полученные посредством регулирования уровня. Один из сигналов Pre\_L и Pre\_R может быть введен в декоррелятор и затем подвергнут операции микширования, выполняемой посредством микшера, тем самым получая сигнал с отрегулированной корреляцией.

Предварительно обработанный стереосигнал понижающего микширования может быть введен в многоканальный кодер. Чтобы предоставить многоканальный вывод, совместимый с информацией позиции объектов и информацией конфигурации воспроизведения, заданной посредством конечного пользователя, не только предварительно обработанный сигнал понижающего микширования, но также основывающаяся на каналах дополнительная информация для выполнения многоканального декодирования обязательна. Далее подробно описывается то, как получить основывающуюся на каналах дополнительную информацию, посредством повторного рассмотрения вышеуказанного примера. Предварительно обработанные сигналы L и R понижающего микширования, которые вводятся в многоканальный декодер, могут быть заданы на основе уравнения (5), как указано посредством уравнения (12):

#### Уравнение 12

$$Eng_{L'} = Eng_{L'_L} + Eng_{R'_L}$$

$$= 0.7 Eng_{Obj1} + 0.3 Eng_{Obj2} + 0.5 Eng_{Obj3} + 0.5 Eng_{Obj4}$$

$$Eng_{R'} = Eng_{L'_R} + Eng_{R'_R}$$

$$= 0.3 Eng_{Obj1} + 0.7 Eng_{Obj2} + 0.5 Eng_{Obj3} + 0.5 Eng_{Obj4}$$

Соотношение, с которым каждый из с первого по четвертый объектного сигнала от ОБЪЕКТ1 до ОБЪЕКТ4 распределяется по FL-, RL-, C-, FR- и RR-каналам, может быть определено так, как указано посредством уравнения (13):

Уравнение 13

$$Eng_{FL} = 0.3 Eng_{Obj1} + 0.1 Eng_{Obj2} + 0.2 Eng_{Obj3} + 0.21 \cdot 100/95 \cdot Eng_{Obj4}$$

$$Eng_{RL} = 0.3 Eng_{Obj1} + 0.1 Eng_{Obj2} + 0.2 Eng_{Obj3} + 0.11 \cdot 100/95 \cdot Eng_{Obj4}$$

$$Eng_C = 0.2 Eng_{Obj1} + 0.2 Eng_{Obj2} + 0.2 Eng_{Obj3} + 0.31 \cdot 100/95 \cdot Eng_{Obj4}$$

$$Eng_{FR} = 0.1 Eng_{Obj1} + 0.3 Eng_{Obj2} + 0.2 Eng_{Obj3} + 0.21 \cdot 100/95 \cdot Eng_{Obj4}$$

$$Eng_{RR} = 0.1 Eng_{Obj1} + 0.3 Eng_{Obj2} + 0.2 Eng_{Obj3} + 0.11 \cdot 100/95 \cdot Eng_{Obj4}$$

Предварительно обработанные сигналы L и R понижающего микширования могут быть расширены до 5.1 каналов посредством MPS, как проиллюстрировано на фиг.27. Ссылаясь на фиг.27, параметры ТТТ-модуля ТТТ0 и ОТТ-модулей ОТТА, ОТТВ и ОТТС, возможно, должны быть вычислены в единицах параметрических полос, даже если параметрические полосы не проиллюстрированы для удобства.

ТТТ-модуль ТТТ0 может быть использован в двух различных режимах: энергетически-ориентированном режиме и режиме прогнозирования. Когда используется в энергетически-ориентированном режиме, ТТТ-модуль ТТТ0 требует двух фрагментов информации CLD. Когда используется в режиме прогнозирования, ТТТ-модуль ТТТ0 требует двух фрагментов информации СРС и фрагмента информации ИСС.

Чтобы вычислить информацию CLD в энергетически-ориентированном режиме, энергетическое соотношение сигналов L", R" и С по фиг.27 может быть вычислено с помощью уравнений (6), (10) и (13). Уровень энергии сигнала L" может быть вычислен так, как указано посредством уравнения (14):

Уравнение 14

$$Eng_{L''} = Eng_{RL} + Eng_{RL} = 0.6 Eng_{Obj1} + 0.2 Eng_{Obj2} + 0.4 Eng_{Obj3} + 0.32 \cdot 100/95 \cdot Eng_{Obj4}$$

$$= 0.6 \cdot \frac{1}{3} \cdot \frac{10^{10}}{1+10^{10}} \cdot \frac{10^{10}}{1+10^{10}} \cdot Eng_L$$

$$+ 0.2 \cdot \frac{2}{5} \cdot \frac{10^{10}}{1+10^{10}} \cdot \frac{1}{1+10^{10}} \cdot Eng_L$$

$$+ 0.4 \cdot \frac{1}{1+10^{10}} \cdot Eng_L$$

$$+ 0.32 \cdot 100/95 \cdot \frac{1}{1+10^{10}} \cdot Eng_R$$

Уравнение (14), возможно, также должно быть использовано для того, чтобы вычислить уровень энергии R" или С. Далее информация CLD, используемая в ТТТ-модуле ТТТ0, может быть вычислена на основе уровней энергии сигналов L", R" и С, как указано посредством уравнения (15):

Уравнение 15

$$TTT_{CLD1} = 10 \log_{10} \left( \frac{Eng_{L^*} + Eng_{R^*}}{Eng_{C^*}} \right)$$

$$TTT_{CLD2} = 10 \log_{10} \left( \frac{Eng_{C^*}}{Eng_{R^*}} \right)$$

Уравнение (14) может быть установлено на основе уравнения (10). Несмотря на то, что уравнения (10) задает только то, как вычислять значения энергии для L-канала, значения энергии для R-канала могут быть вычислены с помощью уравнения (10). Таким образом, значения CLD и ICC третьего и четвертого ОТТ0-модулей могут быть вычислены на основе значений CLD и ICC первого и второго ОТТ-модулей. Тем не менее, это не обязательно может применяться ко всем древовидным структурам, а только к определенным древовидным структурам для декодирования объектных сигналов. Информация, включенная в объектный поток битов, может быть передана в каждый ОТТ-модуль. Альтернативно, информация, включенная в объектный поток битов, может быть передана только в некоторые ОТТ-модули, а информация, указывающая ОТТ-модули, которые не приняли информацию, может быть получена посредством вычисления.

Параметры, такие как информация CLD и ICC, могут быть вычислены для ОТТ-модулей ОТТА, ОТТВ и ОТТС посредством использования вышеупомянутого способа. Эти многоканальные параметры могут быть введены в многоканальный декодер и затем подвергнуты многоканальному декодированию, тем самым получая многоканальный сигнал, который надлежащим образом подготовлен посредством рендеринга согласно информации позиции объекта и информации конфигурации воспроизведения, требуемой посредством конечного пользователя.

Многоканальные параметры могут включать в себя параметр ADG, если уровень объектных сигналов еще не отрегулирован посредством предварительной обработки. Далее подробно описывается вычисление параметра ADG посредством повторного рассмотрения вышеуказанного примера.

Когда матрица рендеринга установлена таким образом, что уровень третьего объектного сигнала может быть увеличен на 10 дБ, что уровень четвертого объектного сигнала может быть уменьшен на 5 дБ, что уровень компонента третьего объектного сигнала в L может быть увеличен на 10 дБ и что уровень компонента четвертого объектного сигнала в L может быть уменьшен на 5 дБ, соотношение  $Ratio_{ADG,L}$  уровней энергии до и после регулирования уровней третьего и четвертого объектных сигналов может быть вычислено с помощью уравнения (16):

Уравнение 16

$$Ratio_{ADG,L} = \frac{Eng_{L\_after}}{Eng_{L\_before}} = \frac{0.7Eng_{Obj1} + 0.3Eng_{Obj2} + 0.5 \cdot 10^{10} \cdot Eng_{Obj3} + 0.5 \cdot 10^{10} \cdot Eng_{Obj4}}{0.7Eng_{Obj1} + 0.3Eng_{Obj2} + 0.5Eng_{Obj3} + 0.5Eng_{Obj4}}$$

Соотношение  $Ratio_{ADG,L}$  может быть определено посредством подстановки уравнения (10) в уравнение (16). Соотношение  $Ratio_{ADG,R}$  для R-канала также может быть вычислено с помощью уравнения (16). Каждое из соотношений  $Ratio_{ADG,L}$  и  $Ratio_{ADG,R}$  представляет варьирование энергии соответствующей параметрической полосы вследствие регулирования уровней объектных сигналов. Таким образом, значения ADG ADG(L) и ADG(R) могут быть вычислены с помощью соотношений  $Ratio_{ADG,L}$  и  $Ratio_{ADG,R}$ , как указано посредством уравнения (17):

Уравнение 17

$$ADG(L) = 10 \log_{10}(\text{Ratio}_{ADG,L})$$

$$ADG(R) = 10 \log_{10}(\text{Ratio}_{ADG,R})$$

5 После того как ADG-параметры ADG(L) и ADG(R) определены, ADG-параметры ADG(L) и ADG(R) квантуются с помощью таблицы квантования ADG, и передаются квантованные значения ADG. Если есть необходимость дополнительно точно отрегулировать ADG-значения ADG(L) и ADG(R), регулирование ADG-значений ADG(L) и ADG(R) может быть выполнено посредством препроцессора, а не  
10 посредством MPS-декодера.

Число и интервал параметрических полос для представления объектных сигналов в объектном потоке битов может отличаться от числа и интервала параметрических полос, используемого в многоканальном декодере. В этом случае параметрические  
15 полосы объектного потока битов могут линейно преобразовываться в параметрические полосы многоканального декодера. Более конкретно, если определенная параметрическая полоса объектного потока битов охватывает две параметрические полосы многоканального декодера, линейное преобразование может быть выполнено таким образом, что определенная параметрическая полоса  
20 объектного потока битов может быть разделена согласно соотношению, с которым соответствующая параметрическая полоса распределена между двумя параметрическими полосами многоканального декодера. С другой стороны, если больше одной параметрической полосы объектного потока битов включено в определенную параметрическую полосу многоканального декодера, значения параметров объектного потока битов могут быть усреднены. Альтернативно, преобразование параметрической полосы может быть выполнено с помощью  
25 существующей таблицы преобразования параметрических полос многоканального стандарта.

30 Когда объектное кодирование используется для телеконференций, голоса различных людей соответствуют объектным сигналам. Объектный декодер выводит голоса, надлежащим образом соответствующие объектным сигналам, в определенные динамики. Тем не менее, когда несколько человек говорят одновременно, объектному  
35 декодеру трудно надлежащим образом распределять голоса людей по различным динамикам посредством декодирования, и рендеринг голосов людей может вызывать звуковые искажения и ухудшать качество звука. Чтобы разрешить это, информация, указывающая то, разговаривают ли несколько человек одновременно, может быть включена в поток битов. Далее, если определено на основе информации, что  
40 несколько человек разговаривают одновременно, основывающийся на каналах поток битов может быть модифицирован так, что едва декодированные сигналы, почти как сигналы понижающего микширования, могут быть выведены в каждый динамик.

Например, предположим, что есть три человека a, b и c, и голоса трех человек a, b и c должны быть декодированы и тем самым выведены в динамики A, B и C  
45 соответственно. Когда три человека a, b и c говорят одновременно, голоса трех человек a, b и c все могут быть включены в сигнал понижающего микширования, который получается посредством понижающего микширования объектных сигналов, соответственно представляющих голоса трех человек a, b и c. В этом случае  
50 информация, касающаяся частей сигнала понижающего микширования, надлежащим образом соответствующих голосам трех человек a, b и c, может быть сконфигурирована как многоканальный поток битов. Далее, сигнал понижающего микширования может быть декодирован с помощью типичного способа

декодирования объектов, с тем, чтобы голоса трех человек а, b и с могли быть выведены в динамики А, В и С соответственно. Вывод каждого из динамиков А, В и С, тем не менее, может быть искаженным и тем самым может иметь меньший коэффициент распознавания, чем исходный сигнал понижающего микширования.

5 Помимо того, голоса трех человек а, b и с могут быть ненадлежащим образом отделены друг от друга. Чтобы разрешить это, информация, указывающая то, что одновременные фрагменты речи трех человек а, b и с произносятся, может быть включена в поток битов. После этого транскодер может сформировать

10 многоканальный поток битов, с тем, чтобы сигнал понижающего микширования, полученный посредством понижающего микширования объектных сигналов, надлежащим образом соответствующих голосам трех человек а, b и с, мог быть выведен в каждый из динамиков А, В и С как есть. Таким образом, можно предотвратить искажения сигнала.

15 В действительности, когда несколько человек говорят одновременно, трудно отделить голос каждого человека. Следовательно, качество звука может быть выше, когда сигнал понижающего микширования выводится как есть, чем когда сигнал понижающего микширования подготавливается посредством рендеринга так, чтобы

20 голоса различных людей могли быть отделены друг от друга и выведены в различные динамики. Для этого транскодер может сформировать многоканальный поток битов, с тем, чтобы сигнал понижающего микширования, полученный из одновременных фрагментов речи нескольких человек, мог быть выведен во все динамики или чтобы сигнал понижающего микширования мог быть усилен и затем выведен в динамики.

25 Чтобы указать то, исходит ли сигнал понижающего микширования из одновременных фрагментов речи нескольких человек, объектный кодер может надлежащим образом модифицировать объектный поток битов вместо предоставления дополнительной информации, как описано выше. В этом случае

30 объектный декодер может выполнить типичную операцию декодирования с объектным потоком битов, с тем, чтобы сигнал понижающего микширования мог быть выведен в динамики как есть или чтобы сигнал понижающего микширования мог быть усилен, но не до такой степени, когда возникают искажения сигнала, и затем выведен в динамики.

35 Далее подробно описывается 3D информация, такая как HTRF, которая предоставляется в многоканальный декодер.

Когда объектный кодер работает в стереофоническом режиме, многоканальный декодер в объектном декодере также работает в стереофоническом режиме. Конечный

40 пользователь может передавать 3D информацию, такую как HRTF, которая оптимизируется на основе пространственных позиций объектных сигналов, в многоканальный декодер.

Более конкретно, когда имеются два объектных сигнала, т.е. ОБЪЕСТ1 и ОБЪЕСТ2, и два объектных сигнала ОБЪЕСТ1 и ОБЪЕСТ2 размещаются в позициях 1 и 2

45 соответственно, формирователь матриц рендеринга или транскодер может иметь 3D информацию, указывающую позиции объектных сигналов ОБЪЕСТ1 и ОБЪЕСТ2. Если формирователь матриц рендеринга имеет 3D информацию, указывающую позиции объектных сигналов ОБЪЕСТ1 и ОБЪЕСТ2, формирователь матриц рендеринга может

50 передавать 3D информацию, указывающую позиции объектных сигналов ОБЪЕСТ1 и ОБЪЕСТ2, в транскодер. С другой стороны, если транскодер имеет 3D информацию, указывающую позиции объектных сигналов ОБЪЕСТ1 и ОБЪЕСТ2, формирователь матриц рендеринга может передавать только индексную информацию,

соответствующую 3D информации, в транскодер.

В этом случае стереофонический сигнал может формироваться на основе 3D информации, задающей позиции 1 и 2, как указано посредством уравнения (18):

**Уравнение 18**

$$L = Obj1 * HRTF_{L,Pos1} + Obj2 * HRTF_{L,Pos2}$$

$$R = Obj1 * HRTF_{R,Pos1} + Obj2 * HRTF_{R,Pos2}$$

Многоканальный стереофонический декодер получает стереофонический звук посредством выполнения декодирования с тем допущением, что 5.1-канальная акустическая система должна быть использована для того, чтобы воспроизводить звук, и стереофонический звук может быть представлен посредством уравнения (19):

**Уравнение 19**

$$L = FL * HRTF_{L,FL} + C * HRTF_{L,C} + FR * HRTF_{L,FR}$$

$$+ RL * HRTF_{L,RL} + RR * HRTF_{L,RR}$$

$$R = FL * HRTF_{R,FL} + C * HRTF_{R,C} + FR * HRTF_{R,FR}$$

$$+ RL * HRTF_{R,RL} + RR * HRTF_{R,RR}$$

Компонент L-канала объектного сигнала ОБЪЕКТ1 может быть представлен посредством уравнения (20):

**Уравнение 20**

$$L_{Obj1} = Obj1 * HRTF_{L,Pos1}$$

$$L_{Obj1} = FL_{Obj1} * HRTF_{L,FL} + C_{Obj1} * HRTF_{L,C} + FR_{Obj1} * HRTF_{L,FR}$$

$$+ RL_{Obj1} * HRTF_{L,RL} + RR_{Obj1} * HRTF_{L,RR}$$

Компонент R-канала объектного сигнала ОБЪЕКТ1 и компоненты L- и R-канала объектного сигнала ОБЪЕКТ2 все могут быть заданы посредством использования уравнения (20). Например, если отношения уровней энергии объектных сигналов ОБЪЕКТ1 и ОБЪЕКТ2 к общему уровню энергии равны a и b соответственно, отношение части объектного сигнала ОБЪЕКТ1, распределяемого в FL-канал, ко всему объектному сигналу ОБЪЕКТ1 равно c, а отношение части объектного сигнала ОБЪЕКТ2, распределяемого в FL-канал, ко всему объектному сигналу ОБЪЕКТ2 равно d, соотношение, с которым объектные сигналы ОБЪЕКТ1 и ОБЪЕКТ2 распределяются в FL-канал, равно ac:bd. В этом случае HRTF FL-канала может быть определено так, как указано посредством уравнения (21):

**Уравнение 21**

$$HRTF_{FL,L} = \frac{ac}{ac+bd} \cdot HRTF_{L,Pos1} + \frac{bd}{ac+bd} \cdot HRTF_{L,Pos2}$$

$$HRTF_{FL,R} = \frac{ac}{ac+bd} \cdot HRTF_{R,Pos1} + \frac{bd}{ac+bd} \cdot HRTF_{R,Pos2}$$

Таким образом, может быть получена 3D информация для использования в многоканальном стереофоническом декодере. Поскольку 3D информация для использования в многоканальном стереофоническом декодере лучше представляет фактические позиции объектных сигналов, можно более живо воспроизводить стереофонические сигналы посредством стереофонического декодирования с помощью 3D информации для применения в многоканальном стереофоническом декодере, чем при выполнении многоканального декодирования с помощью 3D информации, соответствующей пяти позиция динамиков.

Как описано выше, 3D информация для использования в многоканальном стереофоническом декодере может быть основана на 3D информации, представляющей пространственные позиции объектных сигналов и информацию энергетических соотношений. Альтернативно, 3D информация для использования в многоканальном стереофоническом кодере может быть сформирована посредством надлежащего выполнения декорреляции при суммировании 3D информации, представляющей пространственные позиции объектных сигналов, на основе информации ИСС объектных сигналов.

Обработка эффектов может выполняться как часть предварительной обработки. Альтернативно, результат обработки эффектов может просто прибавляться к выводу многоканального декодера. В первом случае для того, чтобы выполнять обработку эффектов объектного сигнала, извлечение объектного сигнала, возможно, должно быть выполнено помимо разделения сигнала L-канала на L\_L и L\_R и разделения сигнала R-канала на R\_R и R\_L.

Более конкретно, объектный сигнал сначала может быть извлечен из сигналов L- и R-канала. Затем сигнал L-канала может быть разделен на L\_L и L\_R, а сигнал R-канала может быть разделен на R\_R и R\_L. Обработка эффектов может выполняться с объектным сигналом. Затем объектный сигнал с обработкой эффектов может быть разделен на компоненты L- и R-канала согласно матрице рендеринга. После этого компонент L-канала объектного сигнала с обработкой эффектов может быть добавлен в L\_L и R\_L, а компонент R-канала объектного сигнала с обработкой эффектов может быть добавлен в R\_R и L\_R.

Альтернативно, могут быть сначала сформированы предварительно обработанные L- и R-сигналы L- и R-канала. Затем объектный сигнал может быть извлечен из предварительно обработанных L- и R-сигналов L- и R-канала. После этого обработка эффектов может быть выполнена для объектного сигнала, и результат обработки эффектов может быть снова добавлен к предварительно обработанным сигналам L- и R-канала.

Спектр объектного сигнала может быть модифицирован посредством обработки эффектов. Например, уровень высокой части или низкой части объектного сигнала может быть выборочно повышен. Для этого только часть спектра, соответствующая высокотоновой части или низкотоновой части объектного сигнала, может быть модифицирована. В этом случае связанная с объектами информация, включенная в объектный поток битов, возможно, должна быть модифицирована соответствующим образом. Например, если уровень низкотоновой части определенного объектного сигнала повышается, энергия низкотоновой части определенного объектного сигнала также может быть повышена. Таким образом, информация энергии, включенная в объектный поток битов, более не представляет корректно энергию определенного объектного сигнала. Чтобы разрешить это, информация энергии, включенная в объектный поток битов, может быть непосредственно модифицирована согласно варьированию энергии определенного объектного сигнала. Альтернативно, информация варьирования спектра, предоставляемая посредством транскодера, может быть применена к формированию многоканального потока битов, с тем, чтобы варьирование в энергии определенного объектного сигнала могло быть отражено в многоканальном битовом потоке.

Фиг.28-33 иллюстрируют схемы для пояснения объединения множества фрагментов основывающейся на объектах дополнительной информации и множества сигналов понижающего микширования во фрагмент дополнительной информации и сигнал

понижающего микширования. В случае телеконференций иногда необходимо комбинировать множество фрагментов основывающейся на объектах дополнительной информации и множества сигналов понижающего микширования в дополнительную информацию и сигнал понижающего микширования. В этом случае

Фиг.28 иллюстрирует схему объектно-кодированного потока битов. Ссылаясь на фиг.28, объектно-кодированный поток битов включает в себя сигнал понижающего микширования и дополнительную информацию. Сигнал понижающего микширования синхронизирован с дополнительной информацией. Следовательно, объектно-кодированный поток битов легко может быть декодирован без учета дополнительных факторов. Тем не менее, в случае объединения множества потоков битов в один поток битов необходимо убедиться в том, что сигнал понижающего микширования одного потока битов синхронизирован с дополнительной информацией одного потока битов.

Фиг.29 иллюстрирует схему для пояснения объединения множества объектно-кодированных потоков битов BS1 и BS2. Ссылаясь на фиг.29, ссылки с номерами 1, 2 и 3 указывают номера кадров. Чтобы объединить множество сигналов понижающего микширования в один сигнал понижающего микширования, сигналы понижающего микширования могут быть преобразованы в сигналы импульсно-кодовой модуляции (PCM), PCM-сигналы могут быть смикшированы с понижением во временной области и смикшированные с понижением PCM-сигналы могут быть преобразованы в формат кодека сжатия. В ходе этих процессов задержка  $d$  может быть сформирована, как проиллюстрировано на фиг.29(b). Следовательно, когда поток битов, который должен быть декодирован, получается посредством объединения множества потоков битов, необходимо убедиться в том, что сигнал понижающего микширования потока битов, который должен быть декодирован, корректно синхронизирован с дополнительной информацией потока битов, который должен быть декодирован.

Если задержка между сигналом понижающего микширования и дополнительной информацией потока битов задана, поток битов может быть скомпенсирован на предварительно определенную величину, соответствующую задержке. Задержка между сигналом понижающего микширования и дополнительной информацией потока битов может варьироваться согласно типу кодека сжатия, используемого для формирования сигнала понижающего микширования. Следовательно, бит, указывающий задержку, если она вообще имеет место, между сигналом понижающего микширования и дополнительной информацией потока битов, может быть включен в дополнительную информацию.

Фиг.30 иллюстрирует объединение двух потоков битов BS1 и BS2 в один поток битов, когда сигналы понижающего микширования потоков битов BS1 и BS2 формируются посредством различных типов кодеков или конфигурация дополнительной информации потока битов BS1 отличается от конфигурации дополнительной информации потока битов BS2. Ссылаясь на фиг.30, когда сигналы понижающего микширования потоков битов BS1 и BS2 формируются посредством различных типов кодеков или конфигурация дополнительной информации потока битов BS1 отличается от конфигурации дополнительной информации потока битов BS2, может быть определено то, что потоки битов BS1 и BS2 имеют различные задержки сигналов  $d1$  и  $d2$ , являющиеся результатом преобразования сигналов понижающего микширования в сигналы временной области и преобразования сигналом временной области с использованием одного кодека сжатия. В этом случае,

если потоки битов BS1 и BS2 просто суммируются без учета различных задержек сигналов, сигнал понижающего микширования потока битов BS1 может быть несогласован с сигналом понижающего микширования потока битов BS2 и дополнительная информация потока битов BS1 может быть несогласованна с 5 дополнительной информацией потока битов BS2. Чтобы разрешить это, сигнал понижающего микширования потока битов BS1, который задерживается на  $d_1$ , может быть дополнительно задержан так, чтобы быть синхронизированным с сигналом понижающего микширования потока битов BS2, который задержан на  $d_2$ . Затем 10 потоки битов BS1 и BS2 могут быть комбинированы с помощью одного способа варианта осуществления по фиг.30. Если имеется более одного потока битов, который должен быть объединен, тот из потоков битов, который имеет наибольшую задержку, может быть использован в качестве опорного потока битов, а другие потоки битов могут быть дополнительно задержаны так, чтобы быть синхронизированными с 15 опорным потоком битов. Бит, указывающий задержку между сигналом понижающего микширования и дополнительной информацией, может быть включен в объектный поток битов.

Бит, указывающий то, есть ли задержка сигнала в потоке битов, может быть 20 предоставлен. Только если информация бита указывает то, что есть задержка сигнала в потоке битов, информация, указывающая задержку сигнала, может быть дополнительно предоставлена. Таким образом, можно минимизировать объем информации, требуемой для указания задержки сигнала, если имеет место, в потоке битов.

Фиг.32 иллюстрирует схему для пояснения того, как компенсировать один из двух потоков битов BS1 и BS2, имеющих различные задержки сигнала, на разность между различными задержками сигнала, в частности, как компенсировать поток битов BS2, который имеет большую задержку сигнала, чем поток битов BS1. Ссылаясь на фиг.32, 30 с первого по третий кадры дополнительной информации потока битов BS1 все могут использоваться как есть. С другой стороны, с первого по третий кадры дополнительной информации потока битов BS2 могут не использоваться как есть, поскольку с первого по третий кадры дополнительной информации потока битов BS2 соответственно не синхронизированы с первым по третий кадрами дополнительной 35 информации потока битов BS1. Например, второй кадр дополнительной информации потока битов BS1 соответствует не только части первого кадра дополнительной информации потока битов BS2, но также части второго кадра дополнительной информации потока битов BS2. Пропорция части второго кадра дополнительной информации потока битов BS2, соответствующей второму кадру дополнительной информации потока битов BS1, ко всему второму кадру дополнительной информации потока битов BS2 и пропорция части первого кадра дополнительной информации потока битов BS2, соответствующей второму кадру дополнительной информации потока битов BS1, ко всему первому кадру дополнительной информации потока 40 битов BS2 может быть вычислена, и первый и второй кадры дополнительной информации потока битов BS2 могут быть усреднены или интерполированы на основе результатов вычисления. Таким образом, с первого по третий кадры дополнительной информации потока битов BS2 соответственно могут быть синхронизированы с 45 первым по третий кадры дополнительной информации потока битов BS1, как проиллюстрировано на фиг.32(b). В таком случае дополнительная информация потока битов BS1 и дополнительная информация потока битов BS2 может быть объединена с помощью способа варианта осуществления по фиг.29. Сигналы понижающего 50

микширования потоков битов BS1 и BS2 могут быть объединены в один сигнал понижающего микширования без необходимости компенсации задержки. В этом случае информация задержки, соответствующая задержке сигнала d1, может быть сохранена в объединенном потоке битов, полученном посредством объединения потоков битов BS1 и BS2.

Фиг.33 иллюстрирует схемы для пояснения того, как компенсировать тот из двух потоков битов, имеющих различные задержки сигнала, который имеет более короткую задержку сигнала. Ссылаясь на фиг.33, с первого по третий кадры дополнительной информации потока битов BS2 все могут использоваться как есть. С другой стороны, с первого по третий кадры дополнительной информации потока битов BS1 могут не использоваться как есть, поскольку с первого по третий кадры дополнительной информации потока битов BS1 соответственно не синхронизированы с первым по третий кадрами дополнительной информации потока битов BS2.

Например, первый кадр дополнительной информации потока битов BS2 соответствует не только части первого кадра дополнительной информации потока битов BS1, но также части второго кадра дополнительной информации потока битов BS1.

Пропорция части первого кадра дополнительной информации потока битов BS1, соответствующей первому кадру дополнительной информации потока битов BS2, ко всему первому кадру дополнительной информации потока битов BS1 и пропорция части второго кадра дополнительной информации потока битов BS1, соответствующей второму кадру дополнительной информации потока битов BS2, ко всему второму кадру дополнительной информации потока битов BS1 может быть вычислена, и первый и второй кадры дополнительной информации потока битов BS1 могут быть усреднены или интерполированы на основе результатов вычисления.

Таким образом, с первого по третий кадры дополнительной информации потока битов BS1 соответственно могут быть синхронизированы с первым по третий кадры дополнительной информации потока битов BS2, как проиллюстрировано на фиг.33(b).

В таком случае дополнительная информация потока битов BS1 и дополнительная информация потока битов BS2 может быть объединена с помощью способа варианта осуществления по фиг.29. Сигналы понижающего микширования потоков битов BS1 и BS2 могут быть объединены в один сигнал понижающего микширования без необходимости компенсации задержки, даже если сигналы понижающего микширования имеют различные задержки сигналов. В этом случае информация задержки, соответствующая задержке сигнала d2, может быть сохранена в объединенном потоке битов, полученном посредством объединения потоков битов BS1 и BS2.

Если множество объектно-кодированных потоков битов объединяются в один поток битов, сигналы понижающего микширования объектно-кодированных потоков битов, возможно, должны быть объединены в один сигнал понижающего микширования. Чтобы объединить множество сигналов понижающего микширования, соответствующих различным кодекам сжатия, в один сигнал понижающего микширования, сигналы понижающего микширования могут быть преобразованы в РСМ-сигналы или сигналы частотной области и РСМ-сигналы или сигналы частотной области могут быть суммированы в соответствующей области. Следовательно, результат суммирования может быть преобразован с помощью предварительно определенного кодека сжатия. Различные задержки сигнала могут возникать согласно тому, сигналы понижающего микширования суммированы в ходе РСМ-операции или суммированы в частотной области, и согласно типу кодека

сжатия. Поскольку декодер не может легко распознавать различные задержки сигнала из потока битов, который должен быть декодирован, информация задержки, касающаяся различных сигналов задержки, возможно, не должна быть включена в поток битов. Эта информация задержки может представлять число выборок задержки в РСМ-сигнале или число выборок задержки в частотной области.

Настоящее изобретение может быть реализовано в качестве машиночитаемого кода, записанного на машиночитаемом носителе записи. Машиночитаемым носителем записи может быть любой тип устройства записи, в котором данные сохраняются машиночитаемым способом. Примеры машиночитаемых носителей записи включают в себя ROM, RAM, CD-ROM, а магнитные ленты, гибкие диски, устройства хранения оптических данных и волновую несущую (например, передачу данных через Интернет). Машиночитаемые носители записи могут распространяться по множеству вычислительных систем, соединенных по сети, так что машиночитаемый код записывается на них и приводится в исполнение с них децентрализованным способом. Функциональные программы, код и сегменты кода, требуемые для реализации настоящего изобретения, могут быть легко истолкованы специалистами в данной области техники.

Как описано выше, согласно настоящему изобретению, звуковые изображения локализуются для каждого объектного сигнала посредством извлечения выгод из преимуществ способов кодирования и декодирования основывающегося на объектах аудио. Таким образом, возможно предлагать более реалистичные звуки в ходе воспроизведения объектных сигналов. Помимо того, настоящее изобретение может быть применено к интерактивным играм и тем самым может предоставлять пользователю более реалистичные переживания в виртуальной реальности.

Несмотря на то, что настоящее изобретение конкретно показано и описано со ссылкой на его примерные варианты осуществления, специалистам в данной области техники следует понимать, что различные изменения по форме и содержанию могут быть сделаны без отступления от духа и объема настоящего изобретения, определяемого прилагаемой формулой изобретения.

#### Формула изобретения

1. Способ декодирования аудио, содержащий этапы, на которых:
  - принимают сигнал понижающего микширования и основывающуюся на объектах дополнительную информацию и управляющую информацию, при этом сигнал понижающего микширования содержит по меньшей мере два канальных сигнала понижающего микширования и сгенерирован посредством понижающего микширования множества объектных сигналов, а управляющая информация управляет позицией или уровнем объектного сигнала, включенного в сигнал понижающего микширования;
  - извлекают из основывающейся на объектах дополнительной информации информацию соотношения распределения между каналами, указывающую соответствующее объектному сигналу соотношение, с которым объектный сигнал распределен по каждому из канальных сигналов понижающего микширования;
  - формируют информацию предварительной обработки для модификации канальных сигналов понижающего микширования на основе информации соотношения распределения между каналами и управляющей информации; и
  - модифицируют канальные сигналы понижающего микширования посредством применения информации предварительной обработки к канальным сигналам

понижающего микширования;

формируют основывающуюся на каналах дополнительную информацию на основе основывающейся на объектах дополнительной информации и управляющей информации; и

5 формируют многоканальный аудиосигнал на основе основывающейся на каналах дополнительной информации и модифицированных канальных сигналов понижающего микширования.

2. Способ декодирования аудио по п.1, в котором основывающаяся на объектах 10 дополнительная информация содержит информацию флага, указывающую то, включена ли информация усиления понижающего микширования в основывающуюся на объектах дополнительную информацию.

3. Способ кодирования аудио, содержащий этапы, на которых: 15 формируют сигнал понижающего микширования посредством понижающего микширования множества объектных сигналов, при этом сигнал понижающего микширования содержит по меньшей мере два канальных сигнала понижающего микширования;

формируют информацию соотношения распределения между каналами 20 посредством вычисления соотношения, с которым объектный сигнал распределяется по каждому из канальных сигналов понижающего микширования;

формируют основывающуюся на объектах дополнительную информацию на основе упомянутого множества объектных сигналов;

25 вставляют информацию соотношения распределения между каналами для модификации канальных сигналов понижающего микширования в основывающуюся на объектах дополнительную информацию; и

формируют поток битов посредством комбинирования сигнала понижающего микширования и основывающейся на объектах дополнительной информации.

30 4. Способ кодирования аудио по п.3, в котором поток битов содержит информацию флага в отношении того, следует ли передавать информацию соотношения распределения между каналами.

5. Устройство декодирования аудио, содержащее:

35 демультимплексор, выполненный с возможностью извлекать сигнал понижающего микширования и основывающуюся на объектах дополнительную информацию из входного аудиосигнала, при этом сигнал понижающего микширования содержит по меньшей мере два канальных сигнала понижающего микширования;

40 преобразователь параметров, выполненный с возможностью извлекать информацию соотношения распределения между каналами, указывающую соответствующее объектному сигналу соотношение, с которым объектный сигнал распределен по каждому из канальных сигналов понижающего микширования, и формировать информацию предварительной обработки для модификации канальных сигналов понижающего микширования на основе информации соотношения

45 распределения между каналами; и

препроцессор, выполненный с возможностью модифицировать канальные сигналы понижающего микширования посредством применения информации предварительной обработки к канальным сигналам понижающего микширования.

50 6. Устройство декодирования аудио по п.5, в котором демультимплексор дополнительно извлекает управляющую информацию, указывающую позицию или уровень объектного сигнала, включенного в сигнал понижающего микширования, при этом преобразователь параметров формирует основывающуюся на каналах

дополнительную информацию на основе основывающейся на объектах дополнительной информации и управляющей информации.

7. Устройство декодирования аудио по п.6, дополнительно содержащее многоканальный кодер, который формирует многоканальный аудиосигнал на основе основывающейся на каналах дополнительной информации и модифицированных канальных сигналов понижающего микширования.

8. Считываемый процессором носитель записи, на который записана программа для выполнения способа по п.1 в процессоре.

9. Считываемый компьютером носитель записи, на котором записана компьютерная программа для выполнения способа декодирования аудио, при этом способ декодирования аудио содержит этапы, на которых:

принимают сигнал понижающего микширования и основывающуюся на объектах дополнительную информацию, при этом сигнал понижающего микширования содержит по меньшей мере два канальных сигнала понижающего микширования;

извлекают из основывающейся на объектах дополнительной информации информацию соотношения распределения между каналами, указывающую соответствующее объектному сигналу соотношение, с которым объектный сигнал распределен по каждому из канальных сигналов понижающего микширования;

формируют информацию предварительной обработки для модификации канальных сигналов понижающего микширования на основе информации соотношения распределения между каналами; и

модифицируют канальные сигналы понижающего микширования посредством применения информации предварительной обработки к канальным сигналам понижающего микширования.

10. Считываемый компьютером носитель записи по п.9, в котором способ декодирования аудио дополнительно содержит этапы, на которых:

принимают управляющую информацию, указывающую позицию или уровень объектного сигнала, включенного в сигнал понижающего микширования;

формируют основывающуюся на каналах дополнительную информацию на основе основывающейся на объектах дополнительной информации и управляющей информации; и

формируют многоканальный аудиосигнал на основе основывающейся на каналах дополнительной информации и модифицированных канальных сигналов понижающего микширования.

11. Считываемый компьютером носитель записи, на котором записана компьютерная программа для выполнения способа кодирования аудио, при этом способ кодирования аудио содержит этапы, на которых:

формируют сигнал понижающего микширования посредством понижающего микширования множества объектных сигналов, при этом сигнал понижающего микширования содержит по меньшей мере два канальных сигнала понижающего микширования;

формируют информацию соотношения распределения между каналами посредством вычисления соотношения, с которым объектный сигнал распределяется по каждому из канальных сигналов понижающего микширования;

формируют основывающуюся на объектах дополнительную информацию на основе упомянутого множества объектных сигналов;

вставляют информацию соотношения распределения между каналами для модификации канальных сигналов понижающего микширования в основывающуюся

на объектах дополнительную информацию; и  
формируют поток битов посредством комбинирования сигнала понижающего  
микширования и основывающейся на объектах дополнительной информации.

5

10

15

20

25

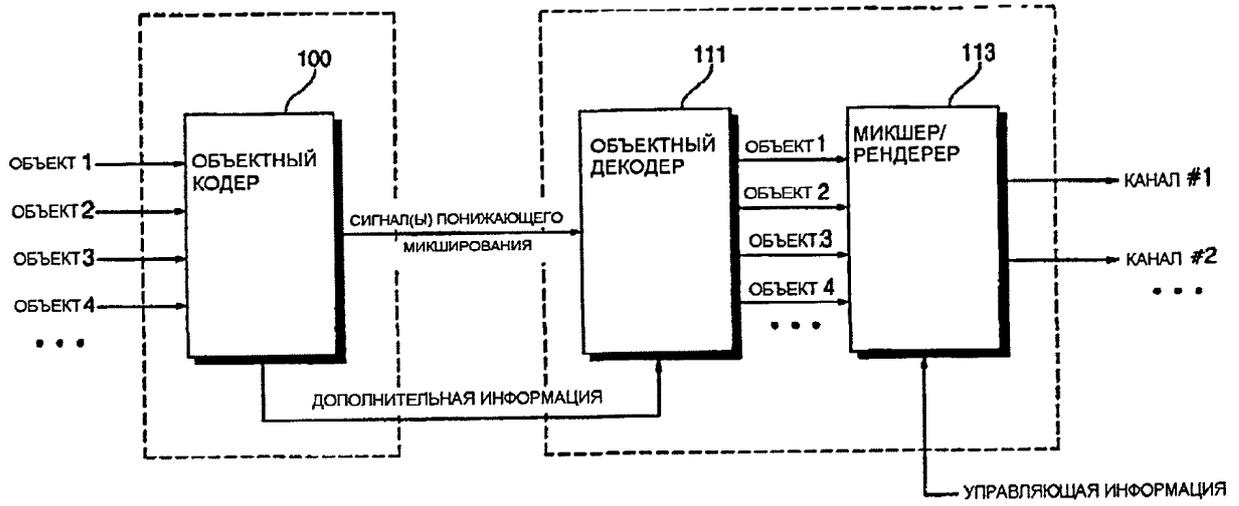
30

35

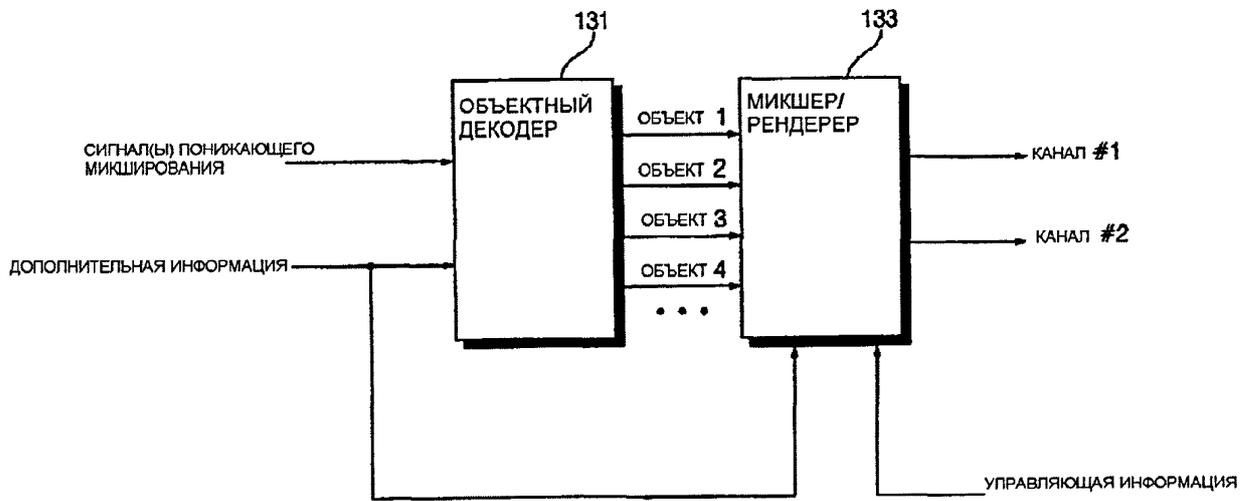
40

45

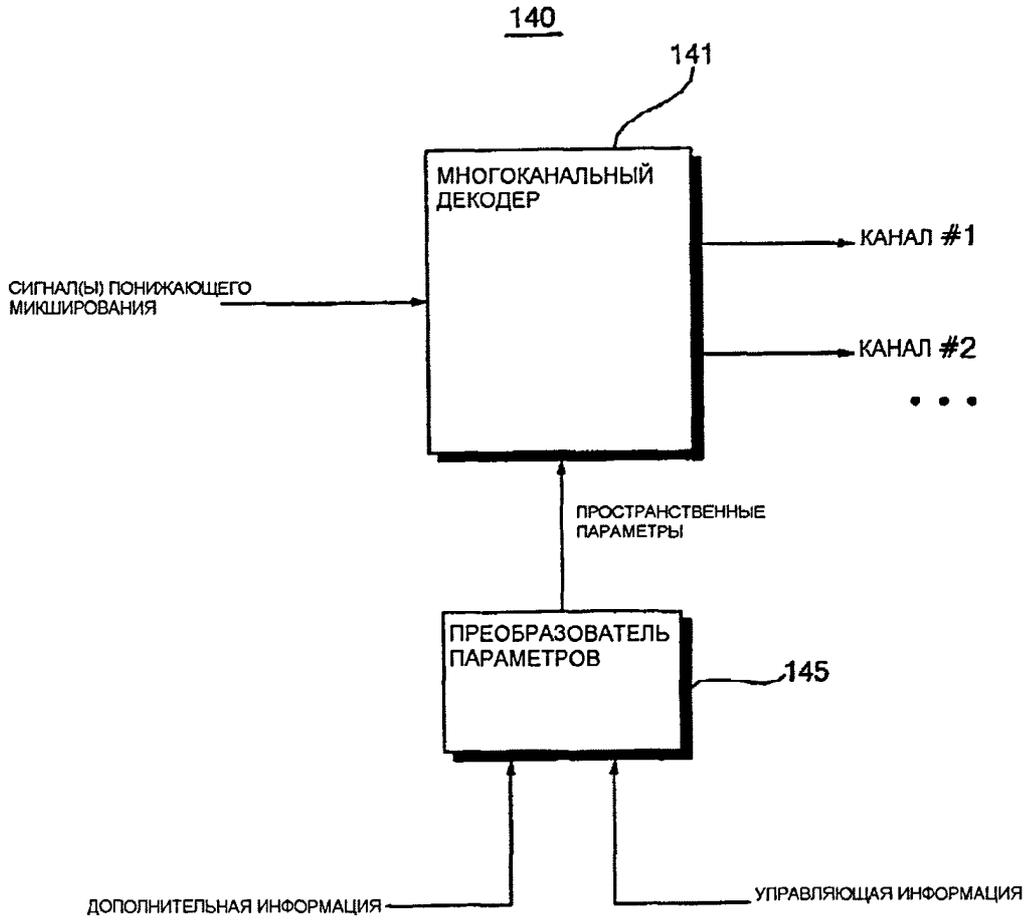
50



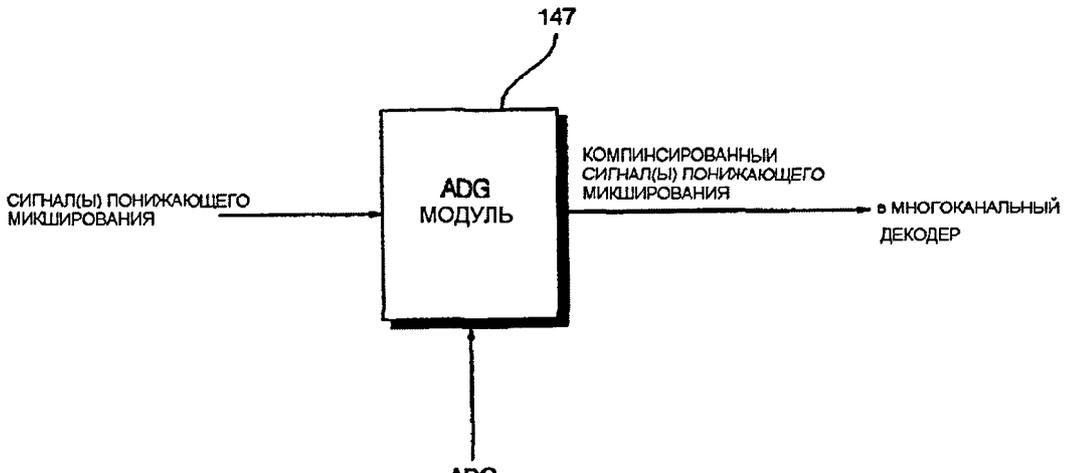
ФИГ.1  
130



ФИГ.3

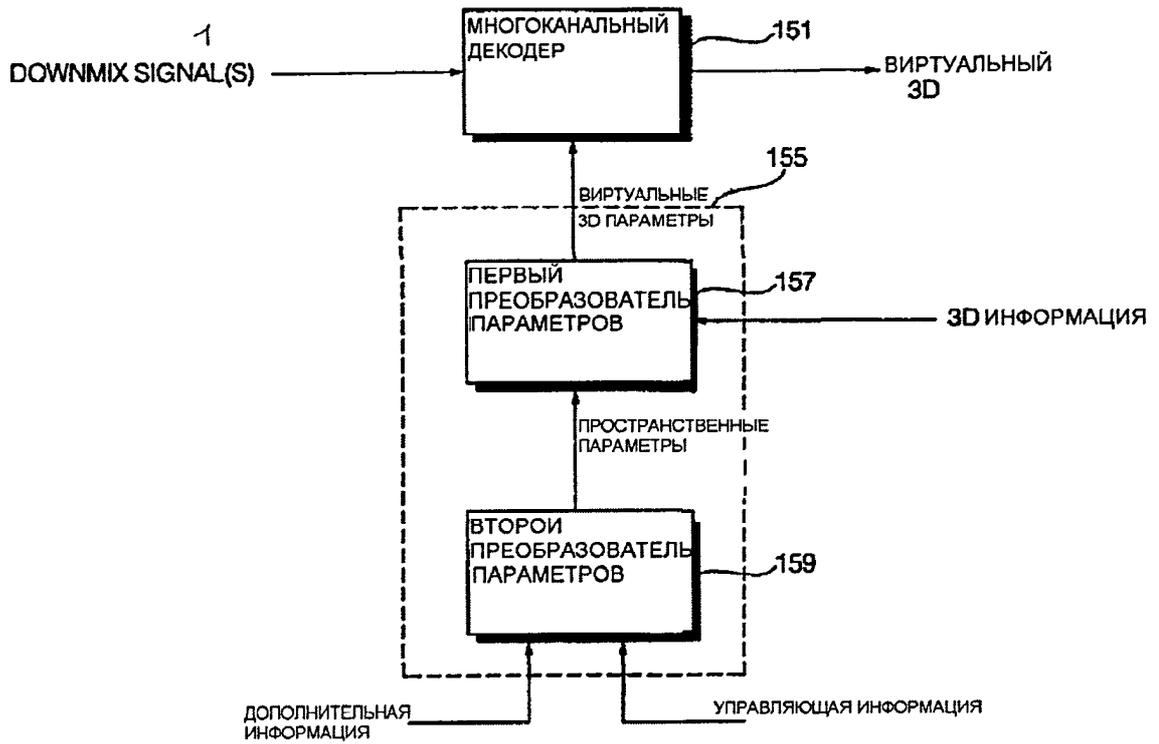


ФИГ.4



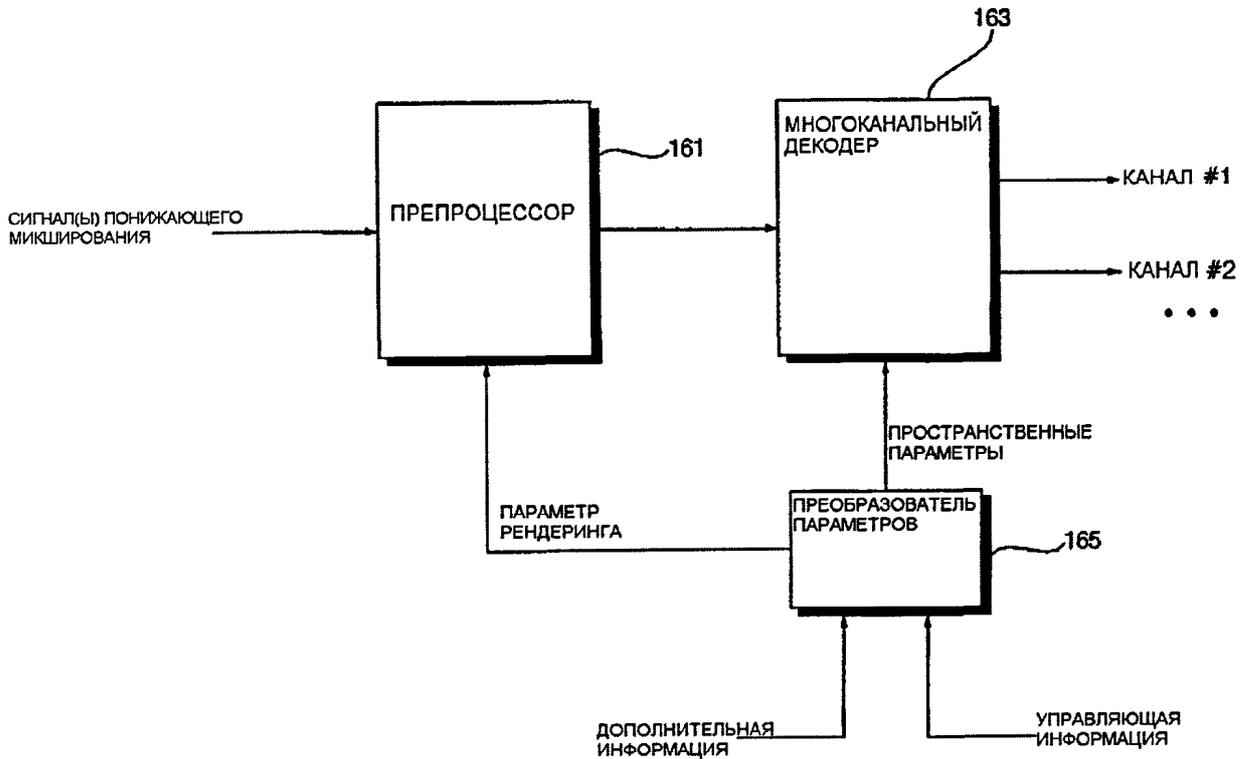
ФИГ.5

150

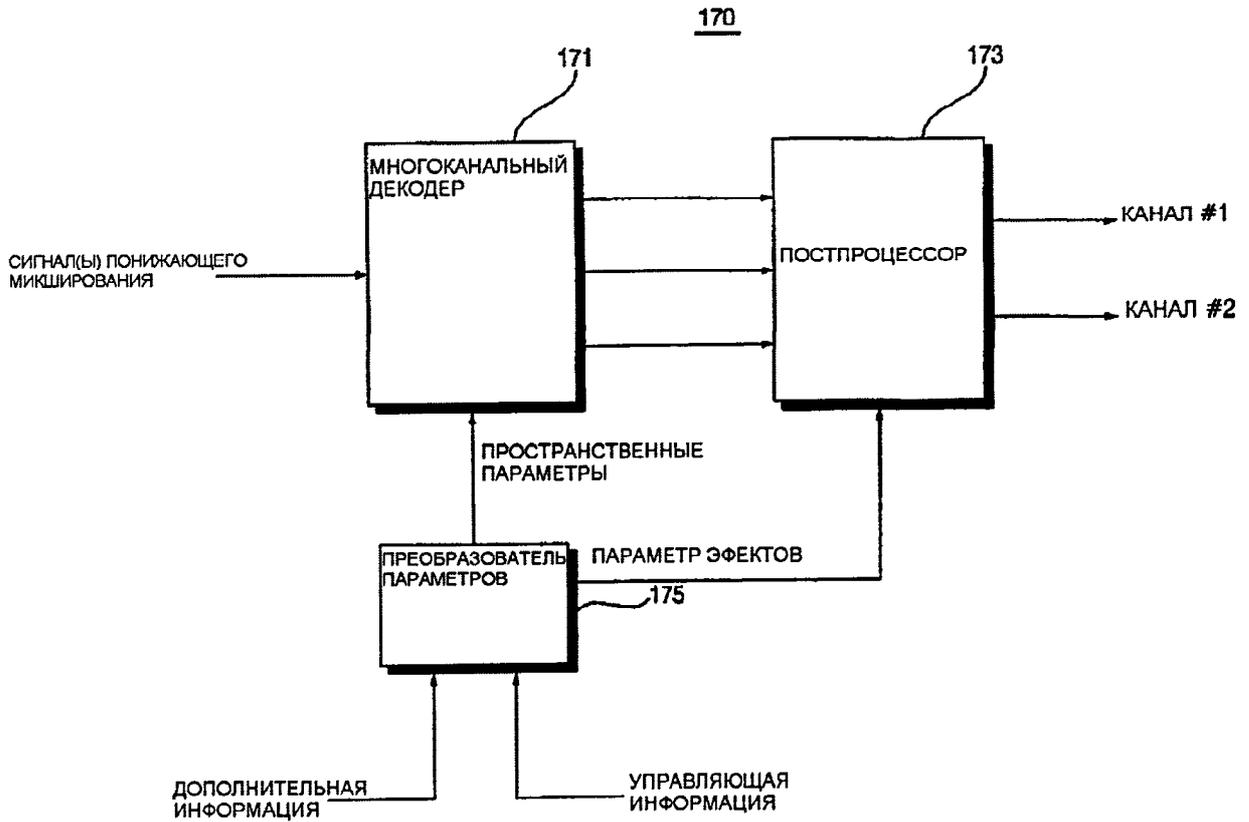


ФИГ.6

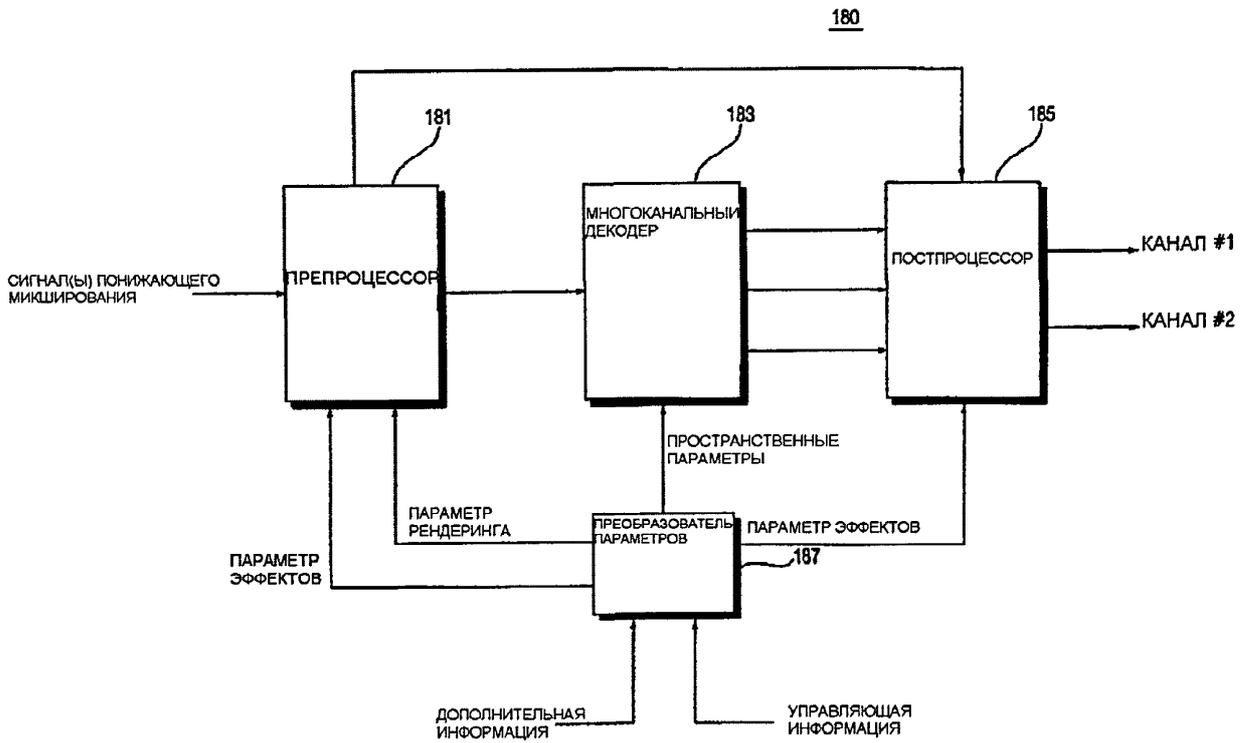
160



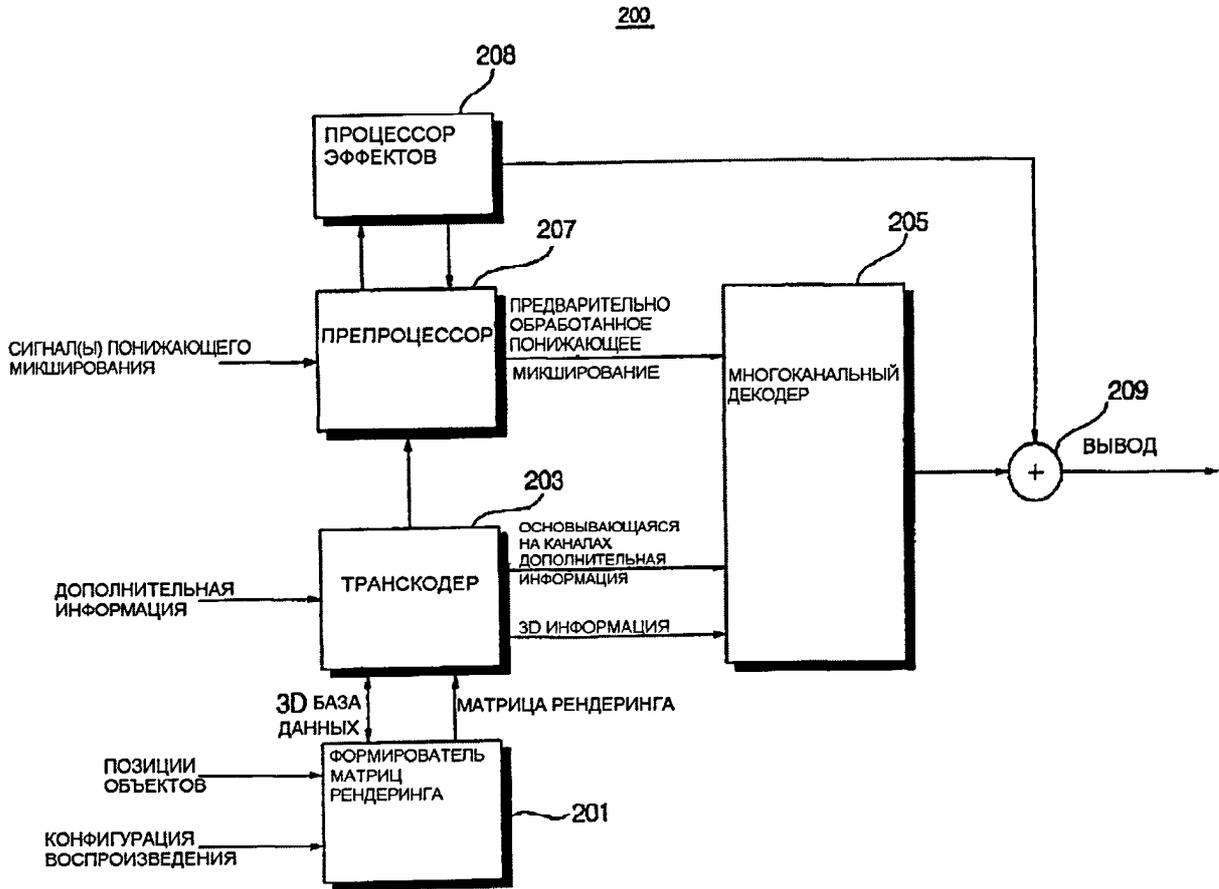
ФИГ.7



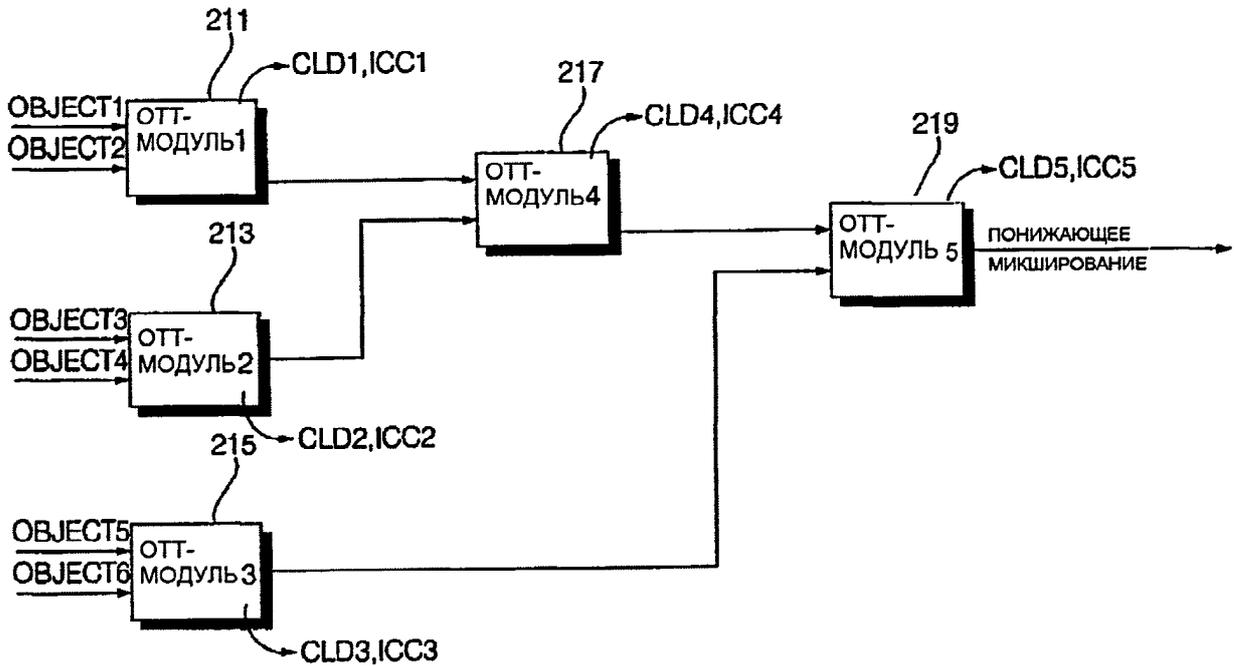
ФИГ.8



ФИГ.9



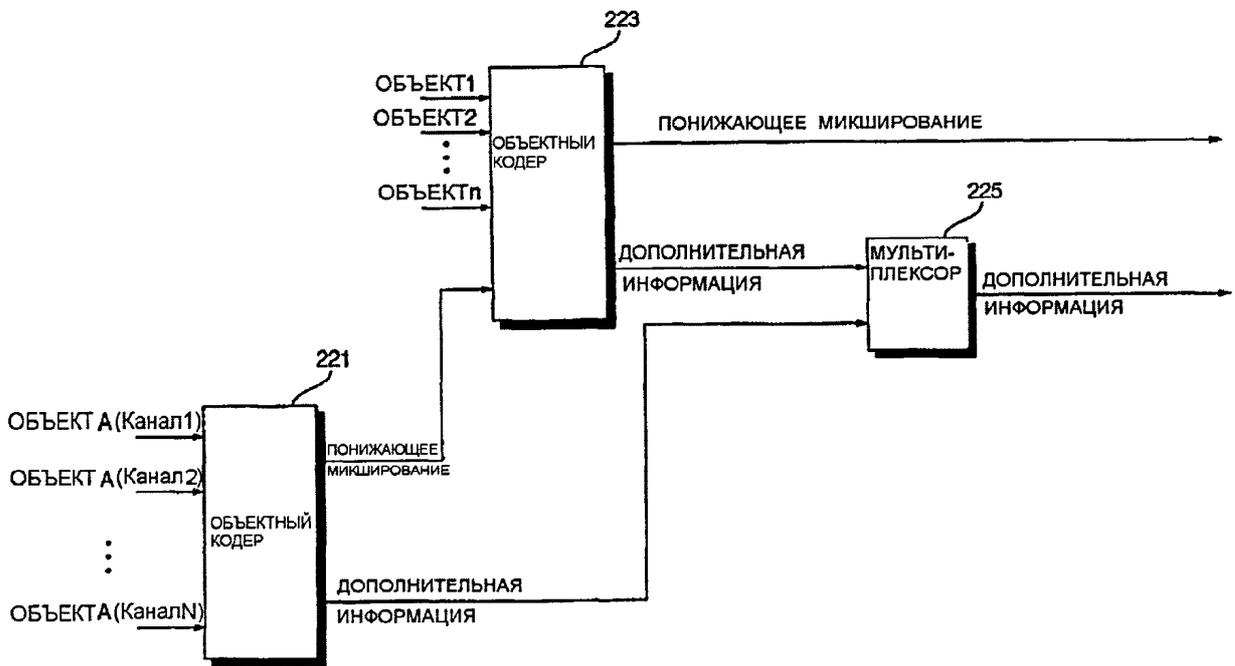
ФИГ.10



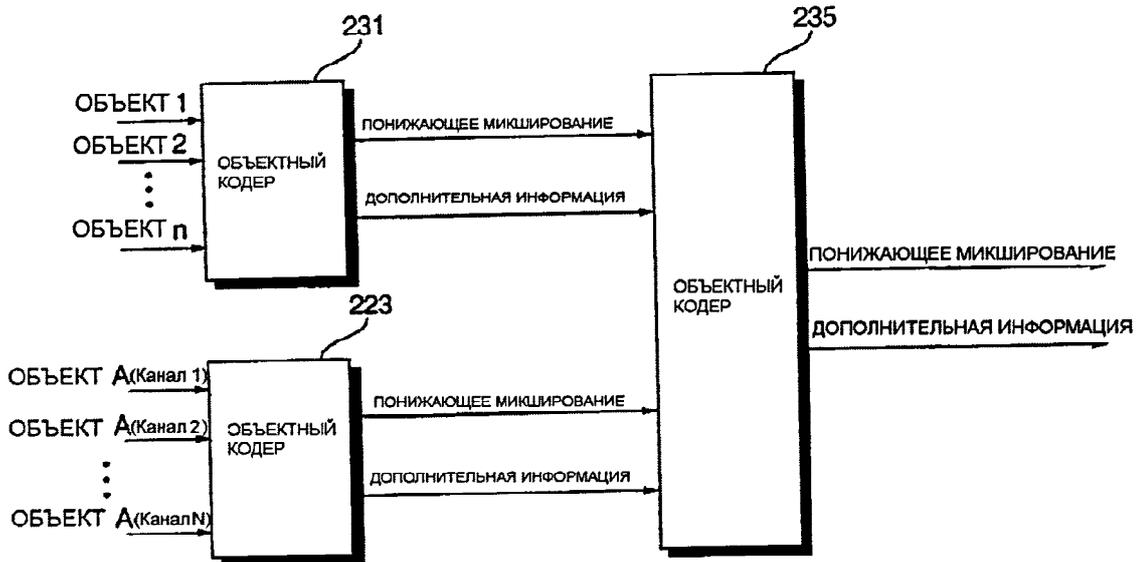
ФИГ.11



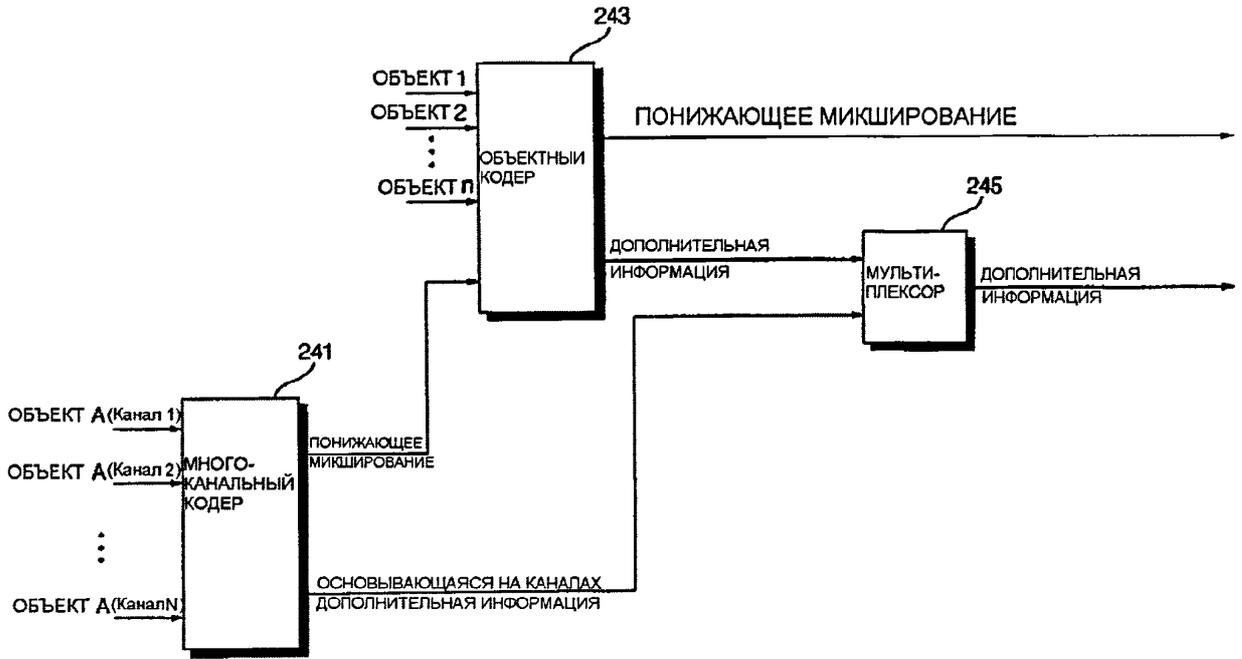
ФИГ.12



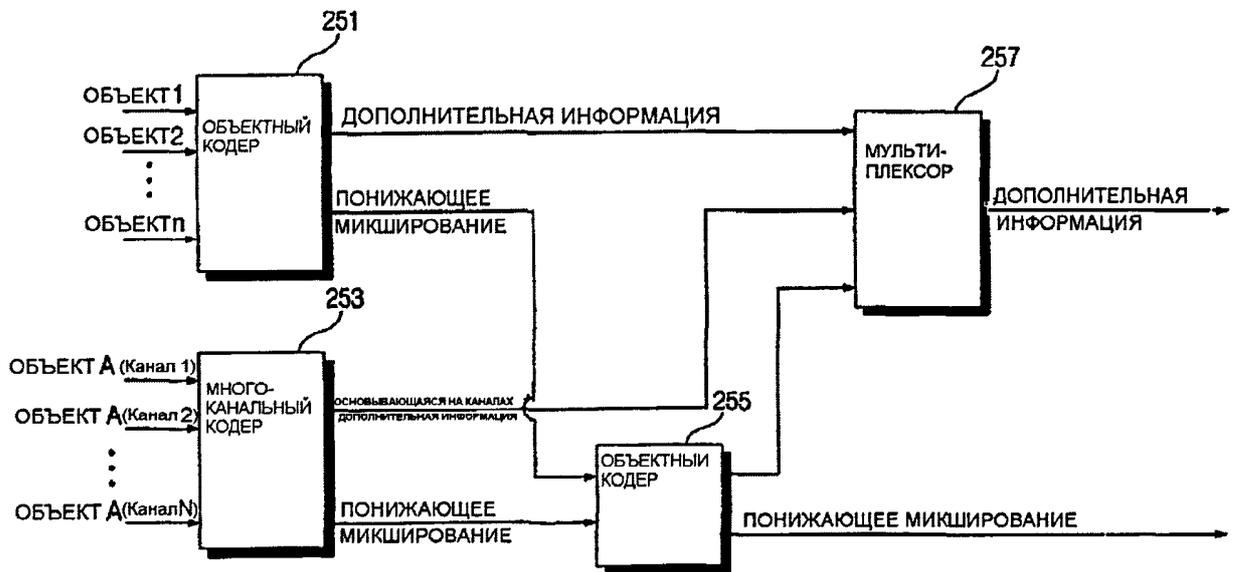
ФИГ.13



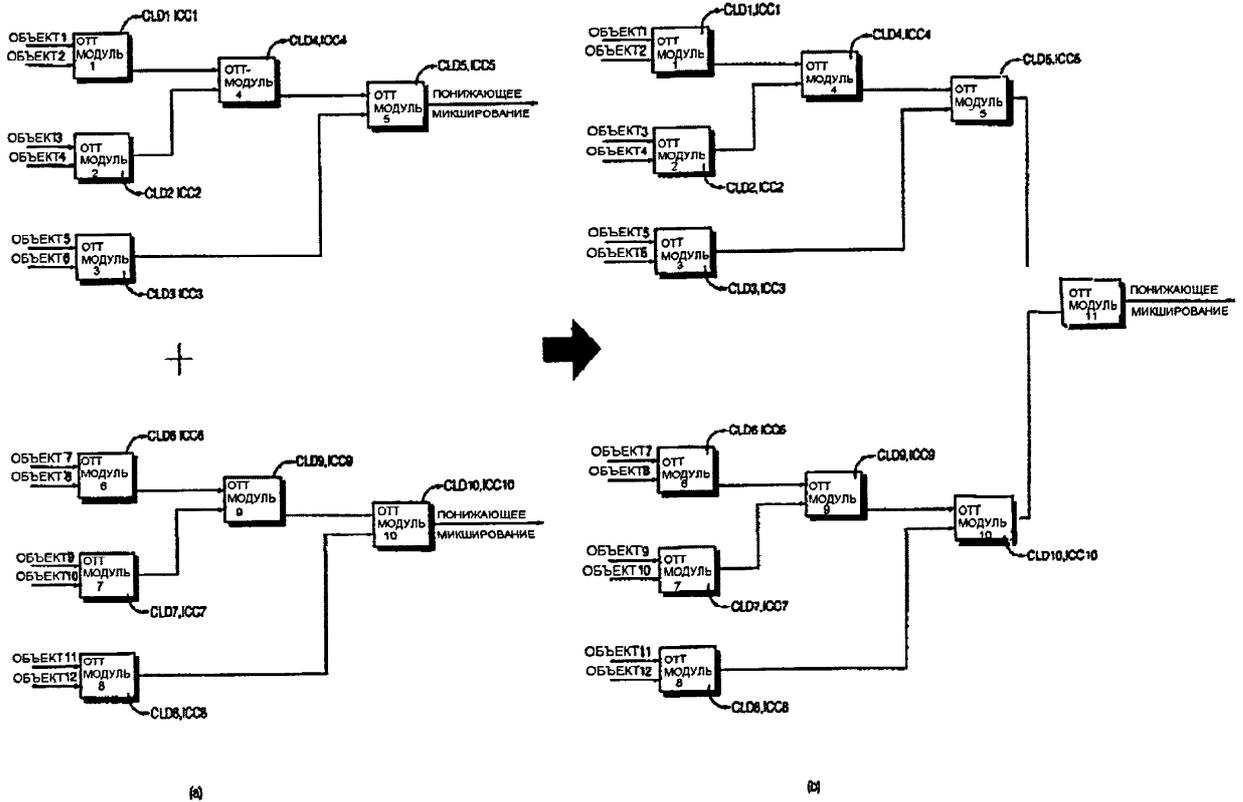
ФИГ.14



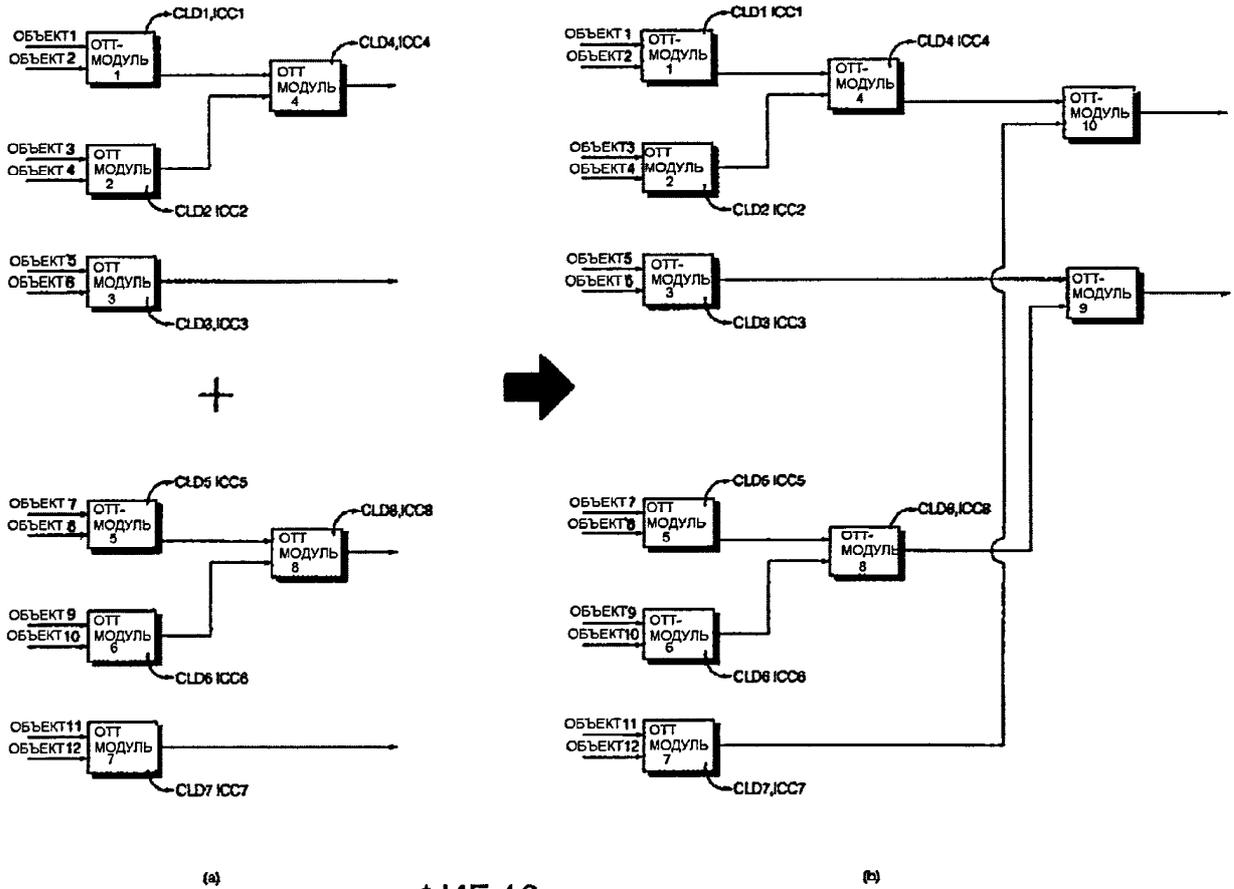
ФИГ.15



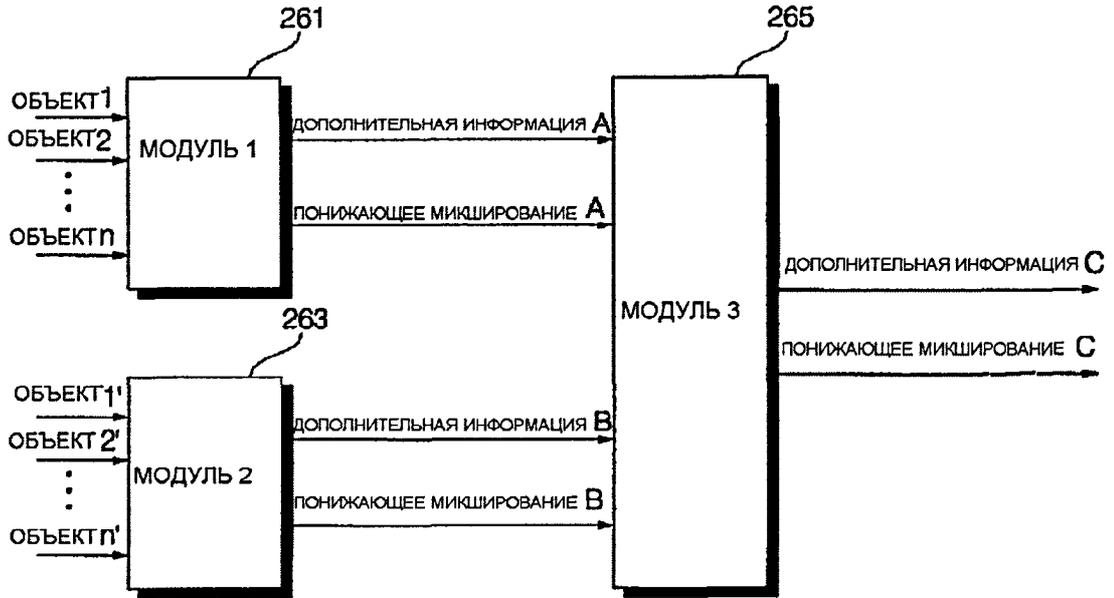
ФИГ.16



ФИГ.17



ФИГ.18

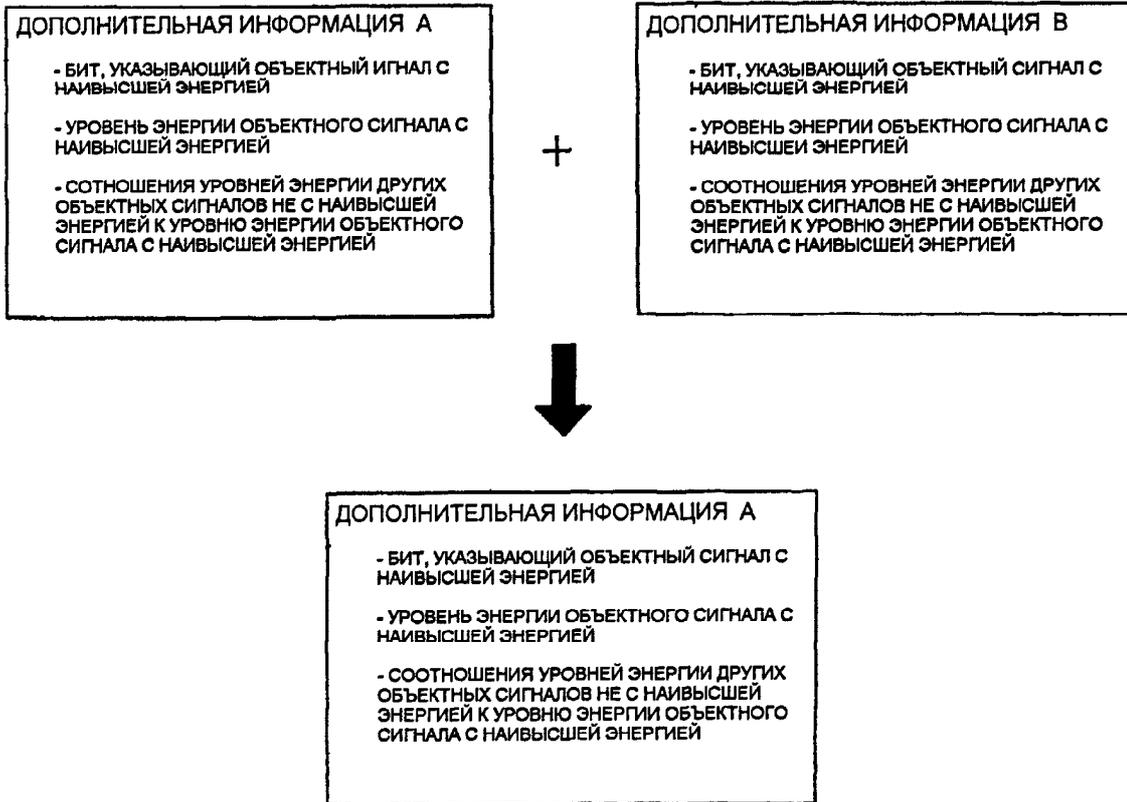


ФИГ.19

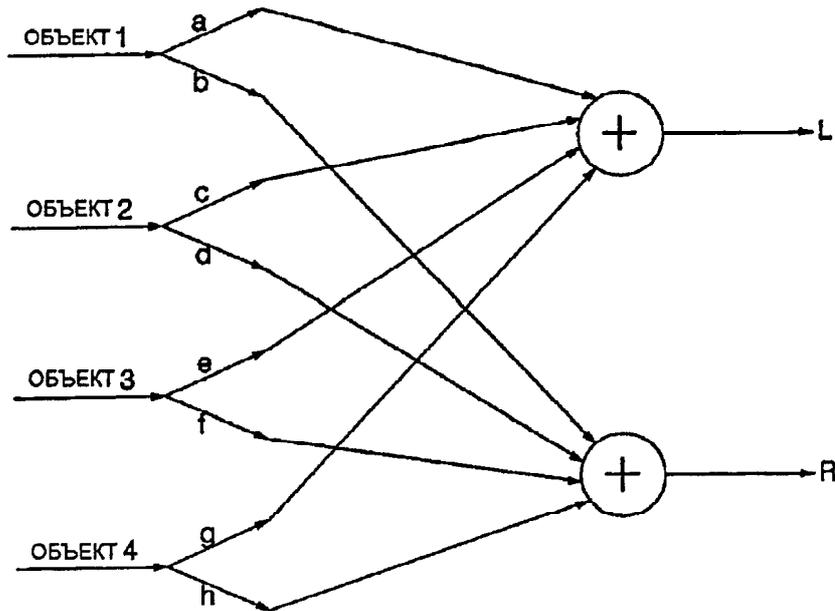


БИТ УКАЗЫВАЮЩИИ ТО ЕСТЬ ЛИ НЕСКОЛЬКО БИТОВЫХ ПОТОКОВ + БИТ УКАЗЫВАЮЩИИ ЧИСЛО БИТОВЫХ ПОТОКОВ	ИДЕНТИФИКАТОР BS	ДОПОЛНИТЕЛЬНАЯ ИНФОРМАЦИЯ A	ИДЕНТИФИКАТОР BS	ДОПОЛНИТЕЛЬНАЯ ИНФОРМАЦИЯ B
---	------------------	--------------------------------	------------------	--------------------------------

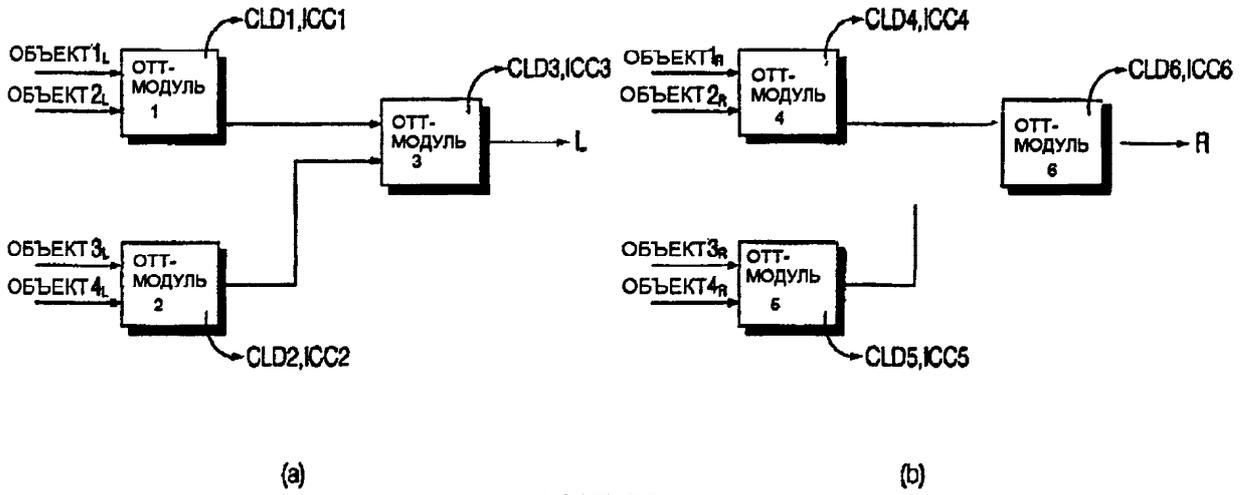
ФИГ.20



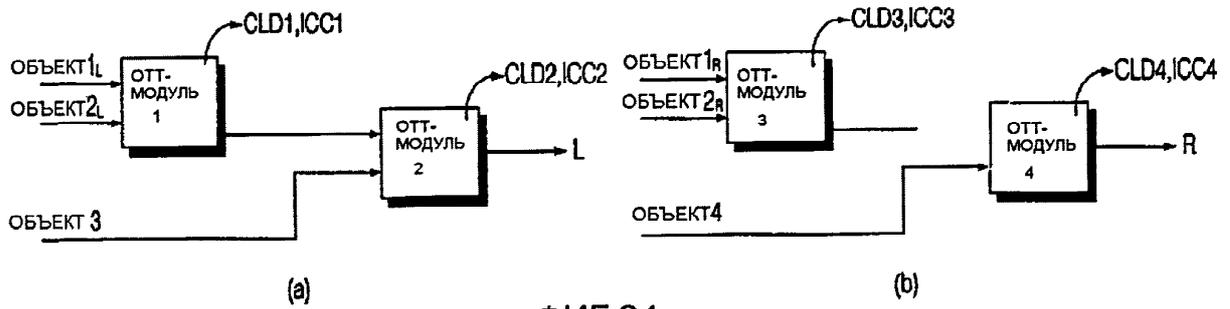
ФИГ.21



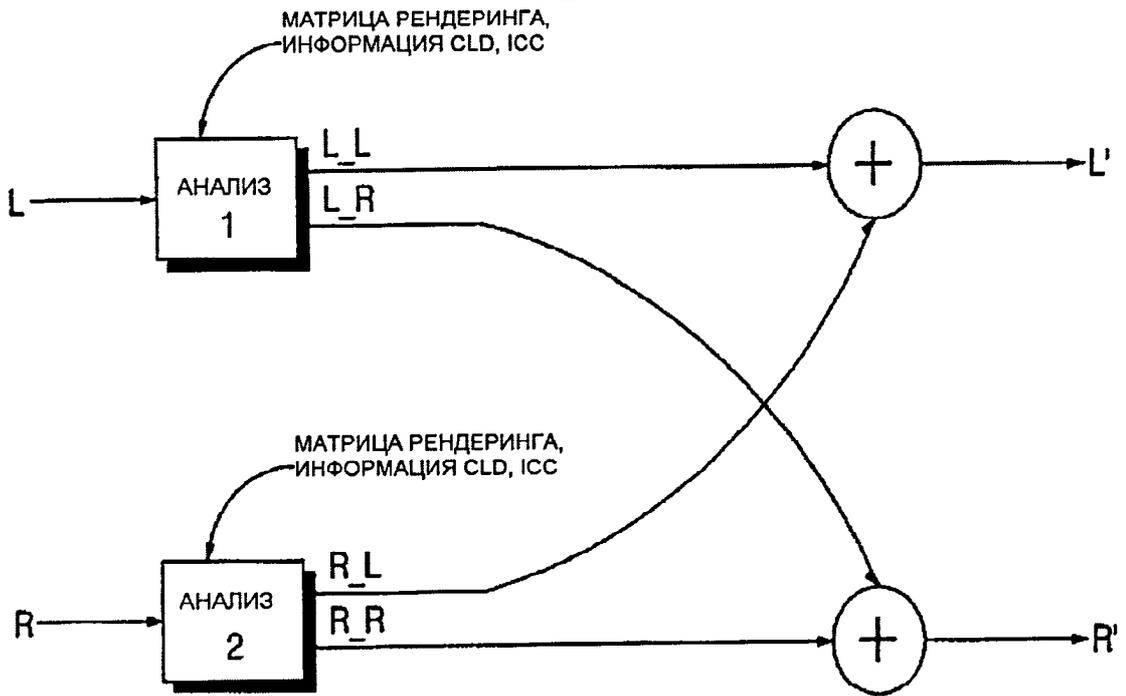
ФИГ.22



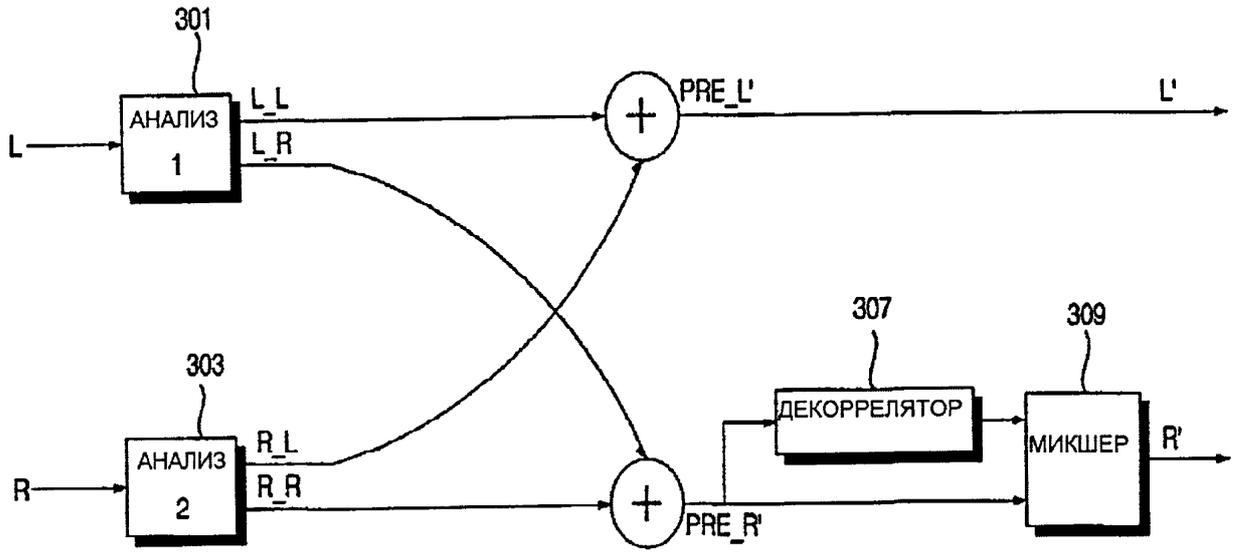
ФИГ.23



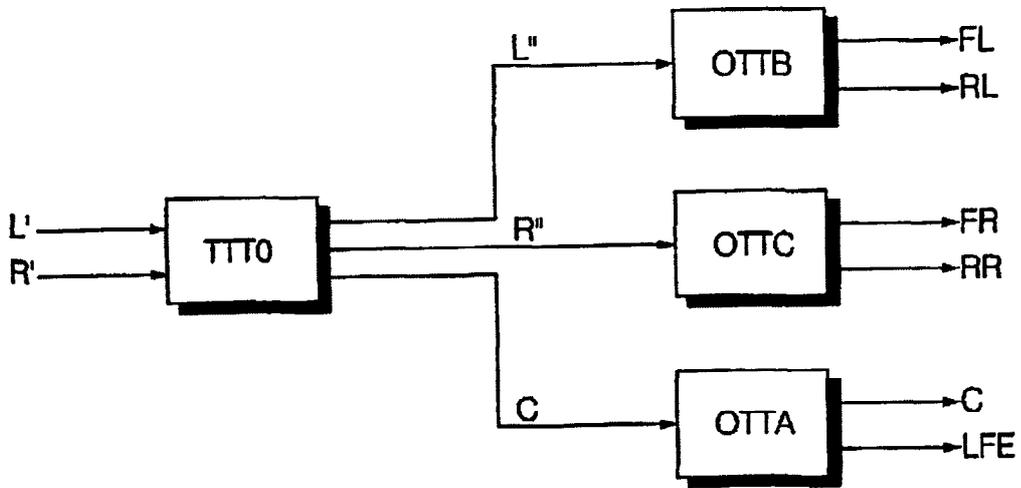
ФИГ.24



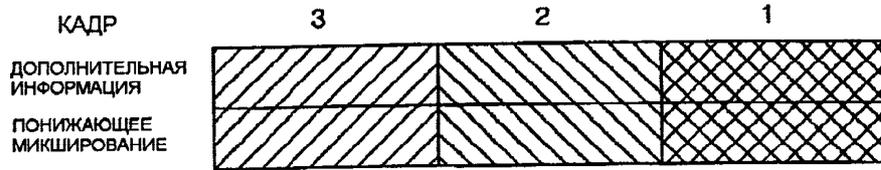
ФИГ.25



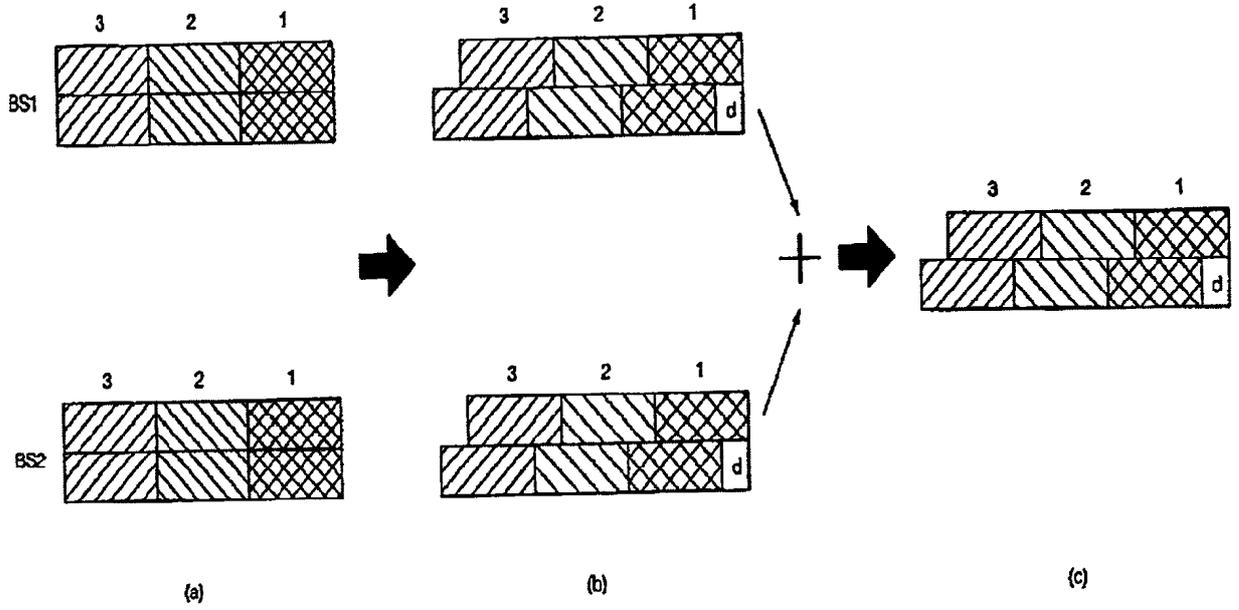
ФИГ.26



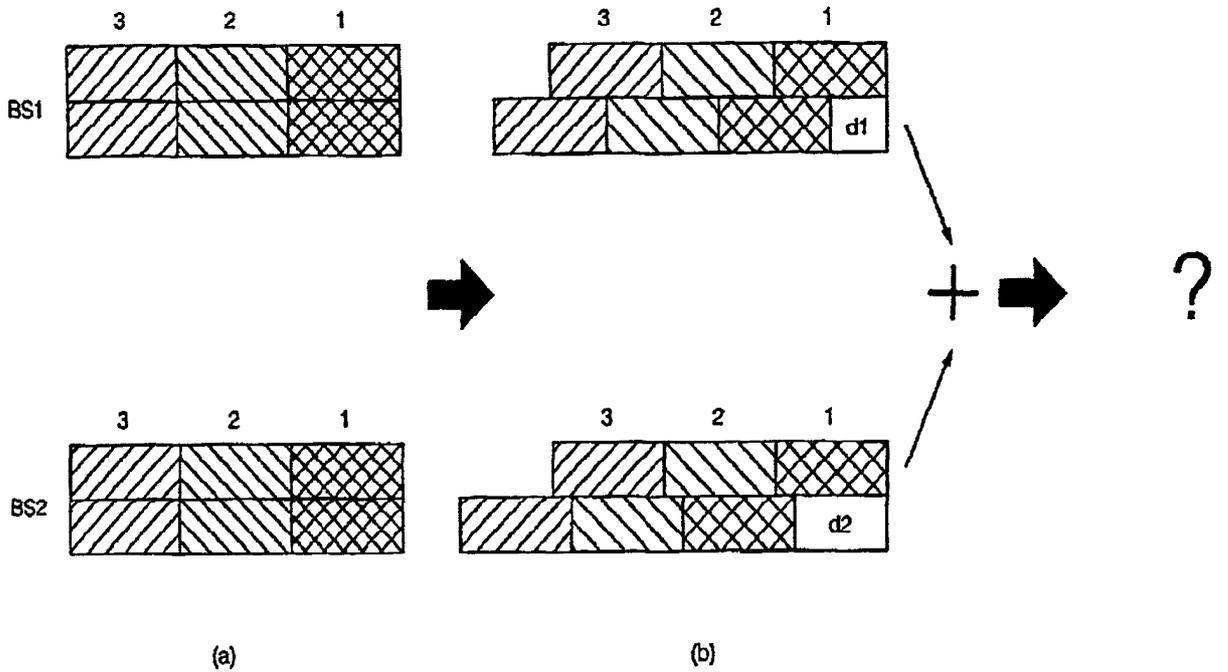
ФИГ.27



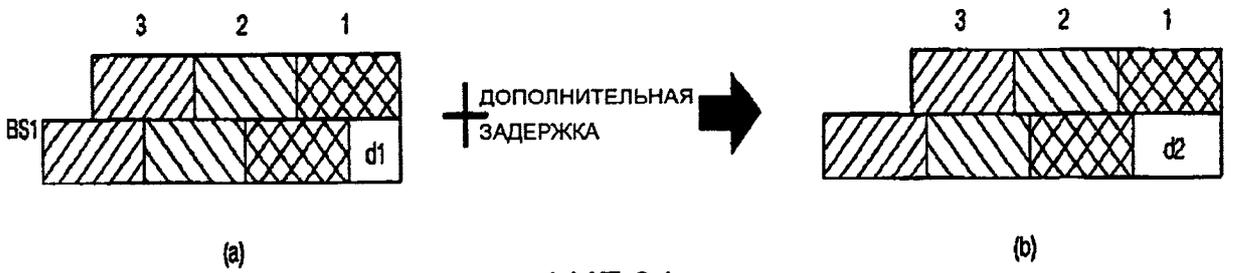
ФИГ.28



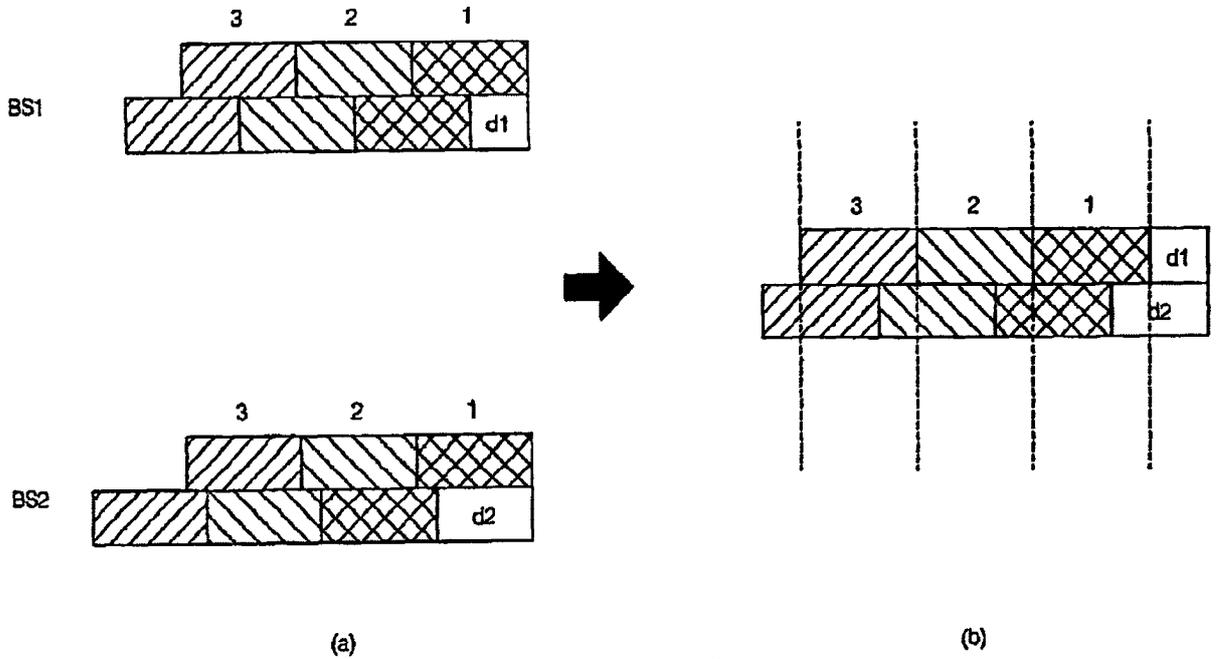
ФИГ.29



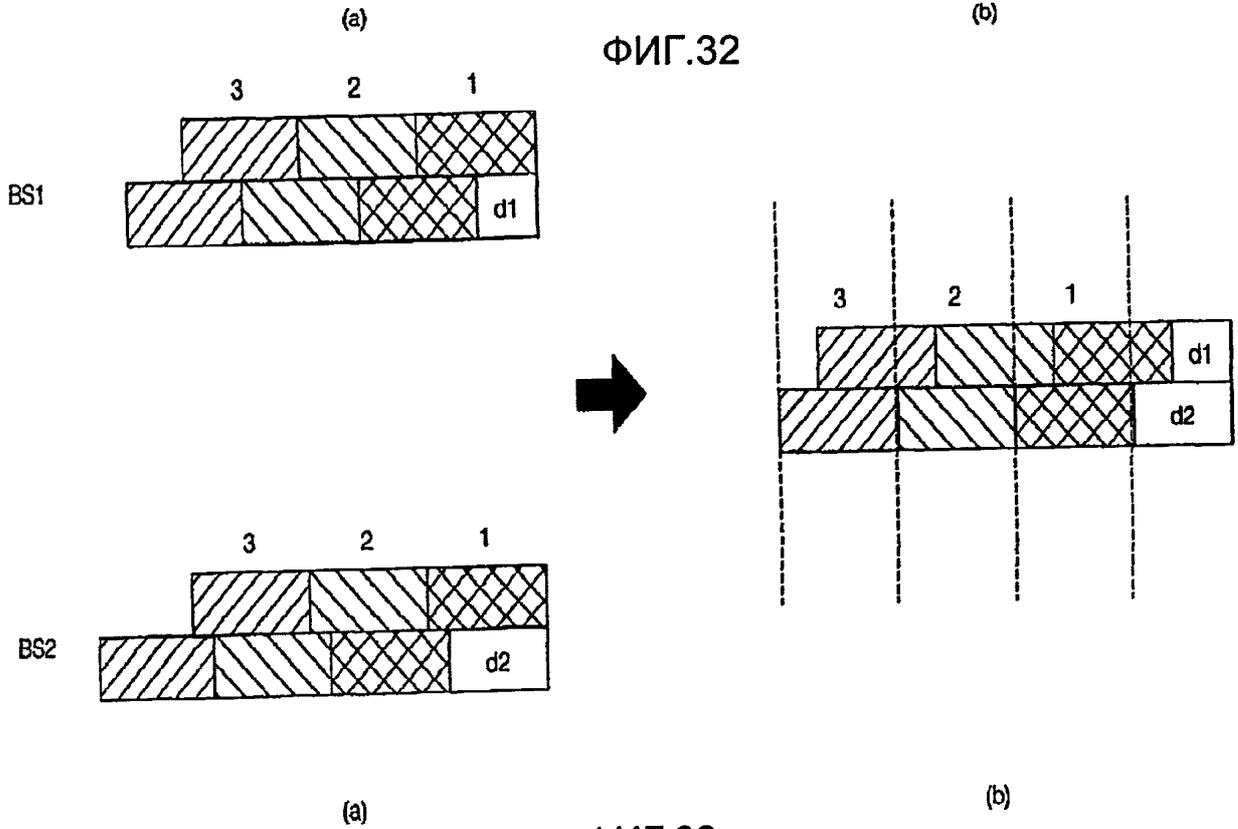
ФИГ.30



ФИГ.31



ФИГ.32



ФИГ.33