



ФЕДЕРАЛЬНАЯ СЛУЖБА  
ПО ИНТЕЛЛЕКТУАЛЬНОЙ СОБСТВЕННОСТИ,  
ПАТЕНТАМ И ТОВАРНЫМ ЗНАКАМ

**(12) ОПИСАНИЕ ИЗОБРЕТЕНИЯ К ПАТЕНТУ**

(21)(22) Заявка: 2008147719/09, 04.05.2007

(24) Дата начала отсчета срока действия патента:  
04.05.2007

Приоритет(ы):

(30) Конвенционный приоритет:

04.05.2006 EP 06113521.6  
13.10.2006 US 60/829,350  
11.01.2007 US 60/884,594  
19.01.2007 US 60/885,742  
06.02.2007 US 60/888,413  
09.03.2007 US 60/894,162

(43) Дата публикации заявки: 10.06.2010 Бюл. № 16

(45) Опубликовано: 10.03.2011 Бюл. № 7

(56) Список документов, цитированных в отчете о  
поиске: EP 1565036 A2, 17.08.2005. RU 2185024  
C2, 10.07.2002. WO 2005/029467 A1, 31.03.2005.  
US 2005195981 A1, 08.09.2005. RU 2129336 C1,  
20.04.1999. US 2005157883 A1, 21.07.2005. WO  
2006008683 A1, 26.01.2006.

(85) Дата начала рассмотрения заявки РСТ на  
национальной фазе: 04.12.2008(86) Заявка РСТ:  
EP 2007/003963 (04.05.2007)(87) Публикация заявки РСТ:  
WO 2007/128523 (15.11.2007)

Адрес для переписки:

129090, Москва, ул. Б.Спаская, 25, стр.3,  
ООО "Юридическая фирма Городиский и  
Партнеры", пат.пов. Ю.Д.Кузнецову,  
рег.№ 595

(72) Автор(ы):

ФАЛЛЕР Кристоф (СН),  
ОХ Хиен О. (KR),  
ДЗУНГ Йанг Вон (KR)

(73) Патентообладатель(и):

ЭлДжи ЭЛЕКТРОНИКС ИНК. (KR)

RU 2 4 1 4 0 9 5 C 2

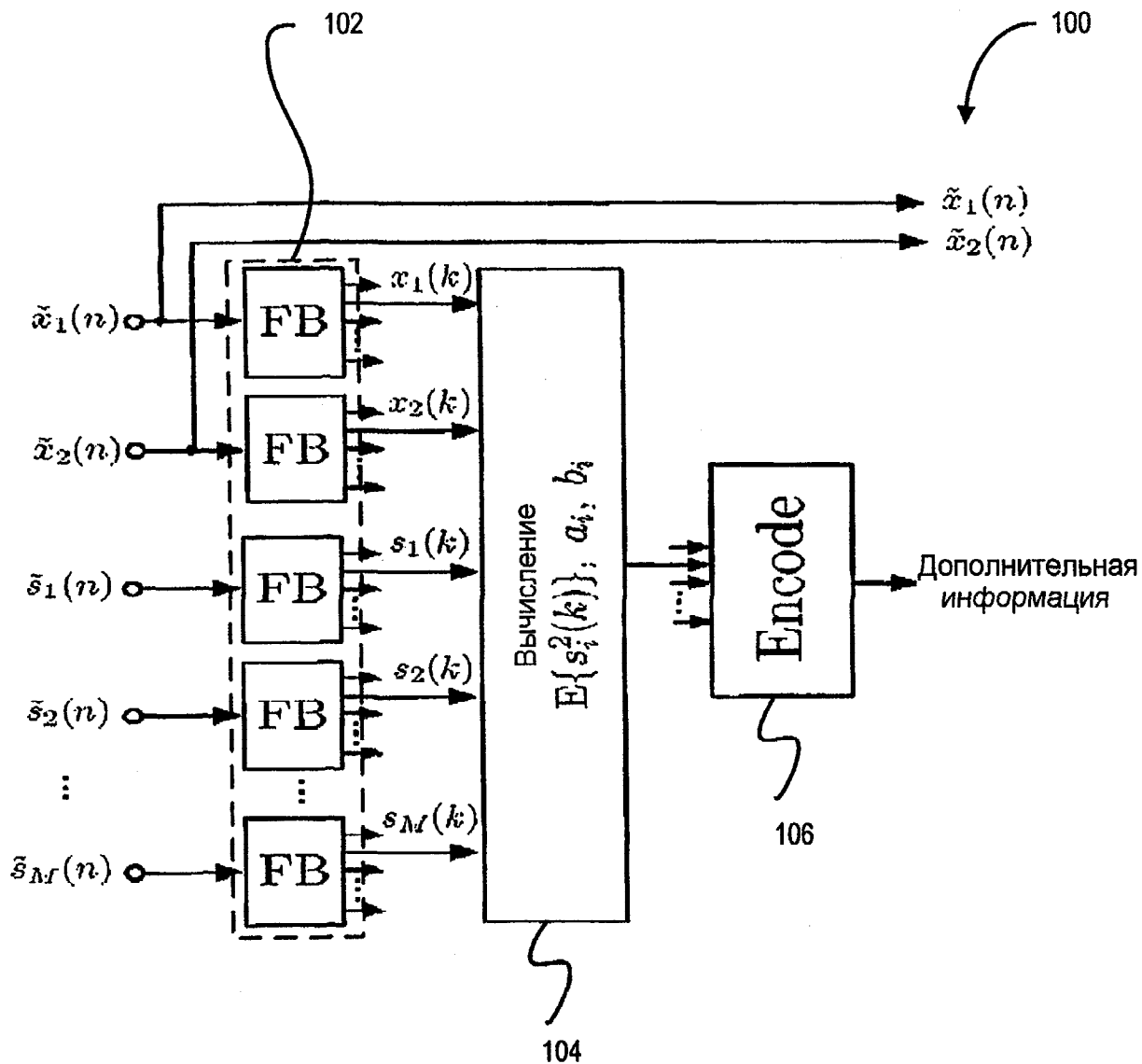
RU 2 4 1 4 0 9 5 C 2

**(54) УСОВЕРШЕНСТВОВАНИЕ ЗВУКОВОГО СИГНАЛА ВОЗМОЖНОСТЬЮ ПОВТОРНОГО МИКШИРОВАНИЯ**

(57) Реферат:

Изобретение относится к обработке звуковых сигналов, Один или более атрибутов (например, панорамирование, усиление и т.п.),

ассоциативно связанных с одним или более объектов (например, инструментов) стереофонического или многоканального звукового сигнала, могут модифицироваться



ФИГ.1А

RU 2414095 C2

RU 2414095 C2



FEDERAL SERVICE  
FOR INTELLECTUAL PROPERTY,  
PATENTS AND TRADEMARKS

**(12) ABSTRACT OF INVENTION**(21)(22) Application: **2008147719/09, 04.05.2007**(24) Effective date for property rights:  
**04.05.2007**

Priority:

(30) Priority:

**04.05.2006 EP 06113521.6**  
**13.10.2006 US 60/829,350**  
**11.01.2007 US 60/884,594**  
**19.01.2007 US 60/885,742**  
**06.02.2007 US 60/888,413**  
**09.03.2007 US 60/894,162**

(43) Application published: **10.06.2010 Bull. 16**(45) Date of publication: **10.03.2011 Bull. 7**(85) Commencement of national phase: **04.12.2008**(86) PCT application:  
**EP 2007/003963 (04.05.2007)**(87) PCT publication:  
**WO 2007/128523 (15.11.2007)**

Mail address:

**129090, Moskva, ul. B.Spaskaja, 25, str.3, OOO**  
**"Juridicheskaja firma Gorodisskij i Partnery",**  
**pat.pov. Ju.D.Kuznetsovu, reg.№ 595**

(72) Inventor(s):

**FALLER Kristof (CH),**  
**OKh Khien O. (KR),**  
**DZUNG Jang Von (KR)**

(73) Proprietor(s):

**EhIDzhi EhLEKTRONIKS INK. (KR)****(54) ENHANCING AUDIO SIGNAL WITH REMIXING CAPABILITY**

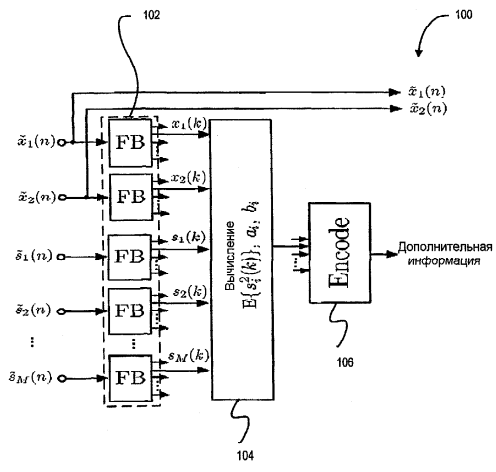
(57) Abstract:

FIELD: physics.

SUBSTANCE: one or more attributes (e.g., pan, gain, etc.) associated with one or more objects (e.g., instruments) of a stereo or multi-channel audio signal can be modified to provide remix capability.

EFFECT: efficient audio signal coding.

50 cl, 26 dwg



ФИГ.1А

## Родственные заявки

Эта заявка испрашивает преимущество приоритета по европейской патентной заявке № EP06113521 на «Enhancing Stereo Audio With Remix Capability» («Усовершенствование стереофонического звукового сигнала возможностью повторного микширования»), поданной 4 мая 2006 года, каковая заявка включена в материалы настоящей заявки посредством ссылки во всей своей полноте.

Эта заявка испрашивает приоритет по предварительной заявке на выдачу патента США под № 60/829,350, на «Enhancing Stereo Audio With Remix Capability», поданной 13 октября 2006 года, каковая заявка включена в материалы настоящей заявки посредством ссылки во всей своей полноте.

Эта заявка испрашивает приоритет по предварительной заявке на выдачу патента США под № 60/884,594 на «Separate Dialogue Volume» («Раздельный уровень громкости диалога»), поданной 11 января 2007 года, каковая заявка включена в материалы настоящей заявки посредством ссылки во всей своей полноте.

Эта заявка испрашивает приоритет по предварительной заявке на выдачу патента США под № 60/885,742 на «Enhancing Stereo Audio With Remix Capability», поданной 19 января 2007 года, каковая заявка включена в материалы настоящей заявки посредством ссылки во всей своей полноте.

Эта заявка испрашивает приоритет по предварительной заявке на выдачу патента США под № 60/888,413 на «Object-Based Signal Reproduction» («Воспроизведение объектно-ориентированного сигнала»), поданной 6 февраля 2007 года, каковая заявка включена в материалы настоящей заявки посредством ссылки во всей своей полноте.

Эта заявка испрашивает приоритет по предварительной заявке на выдачу патента США под № 60/894,162 на «Bitstream and Side Information For SAOC/Remix» («Битовый поток и дополнительная информация для SAOC/повторного микширования»), поданной 9 марта 2007 года, каковая заявка включена в материалы настоящей заявки посредством ссылки во всей своей полноте.

## Область техники

Объект изобретения этой заявки в целом имеет отношение к обработке звуковых сигналов.

## Уровень техники

Многие потребительские звуковые устройства (например, стереосистемы, медиаплееры, мобильные телефоны, игровые консоли и т.п.) предоставляют пользователям возможность модифицировать стереофонические звуковые сигналы с использованием средств управления для коррекции (например, нижних звуковых частот, верхних звуковых частот), уровня громкости, акустических эффектов помещений и т.п. Эти модификации, однако, применяются ко взятому в целом звуковому сигналу, а не к отдельным звуковым объектам (например, инструментам), которые составляют звуковой сигнал. Например, пользователь не может модифицировать по отдельности стереофоническое панорамирование или коэффициент усиления гитар, барабанов или вокальных партий в песне, не воздействуя на песню, взятую в целом.

Были предложены технологии, которые обеспечивают гибкость микширования в декодере. Эти технологии полагаются на кодирование бинауральными контрольными сигналами (ВСС), параметрический или пространственный звуковой декодер для формирования микшированного выходного сигнала декодера. Ни одна из этих технологий, однако, не кодирует непосредственно стереофонические микшированные сигналы (например, профессионально микшированную музыку) для предоставления

возможности обратной совместимости без дискредитации качества звучания.

Технологии пространственного звукового кодирования были предложены для представления стереофонических или многоканальных звуковых каналов с использованием межканальных контрольных сигналов (например, перепада уровней, 5 разновременности, разности фаз, когерентности). Межканальные контрольные сигналы передаются в качестве «дополнительной информации» в декодер для использования при формировании многоканального выходного сигнала. Эти традиционные технологии пространственного звукового кодирования, однако, имеют несколько недостатков. Например, по меньшей мере, некоторые из этих технологий требуют, чтобы отдельный сигнал для каждого звукового объекта передавался в декодер, даже если звуковой объект не будет модифицироваться в декодере. Такое требование имеет следствием излишнюю обработку в кодировщике и декодере. Еще одним недостатком является ограничение выходного сигнала декодера либо 10 стереофоническим (или многоканальным) звуковым сигналом либо звуковым сигналом источника, имея следствием пониженную гибкость для повторного микширования в декодере. В заключение, по меньшей мере, некоторые из этих традиционных технологий требуют сложной декорреляционной обработки в декодере, делающей такие технологии непригодными для некоторых применений или устройств. 20

#### Сущность изобретения

Один или более атрибутов (например, панорамирование, усиление, и т.п.), ассоциативно связанных с одним или более объектов (например, инструментов) стереофонического или многоканального звукового сигнала, могут 25 модифицироваться для предоставления возможности повторного микширования.

В некоторых реализациях способ включает в себя получение первого многоканального звукового сигнала, содержащего набор объектов; получение дополнительной информации, по меньшей мере, некоторая часть которой 30 представляет зависимость между первым многоканальным звуковым сигналом и одним или более сигналами источников, представляющими объекты, которые должны повторно микшироваться; получение набора параметров микширования; и формирование второго многоканального звукового сигнала с использованием дополнительной информации и набора параметров микширования. 35

В некоторых реализациях, способ включает в себя получение звукового сигнала, содержащего набор объектов; получение подмножества сигналов источников, представляющих подмножество объектов; и формирование дополнительной информации из подмножества сигналов источников, по меньшей мере, некоторая 40 часть дополнительной информации представляет взаимосвязь между звуковым сигналом и подмножеством сигналов источников.

В некоторых реализациях, способ включает в себя получение многоканального звукового сигнала; определение коэффициентов усиления для набора сигналов источников с использованием требуемых перепадов уровней источников, 45 представляющих требуемые направления звучания набора сигналов источников в павильоне звукозаписи; оценивание мощности поддиапазона для направления прямого звучания набора сигналов источников с использованием многоканального звукового сигнала; и оценивание мощности поддиапазонов для по меньшей мере 50 некоторых из сигналов источников в наборе сигналов источников посредством модифицирования мощности поддиапазона для направления прямого звучания в качестве функции направления прямого звучания и требуемого направления звучания.

В некоторых реализациях, способ включает в себя получение микшированного

звукового сигнала; получение набора параметров микширования для повторного микширования микшированного звукового сигнала; если дополнительная информация доступна, повторное микширование микшированного звукового сигнала с использованием дополнительной информации и набора параметров микширования; 5  
если дополнительная информация не доступна, формирование набора слепых параметров из микшированного звукового сигнала; и формирование повторно микшированного звукового сигнала с использованием слепых параметров и набора параметров микширования.

10 В некоторых реализациях, способ включает в себя получение микшированного звукового сигнала, включающего в себя сигналы речевого источника; получение параметров микширования, задающих требуемое усовершенствование одному или более речевых сигналов источников; формирование набора слепых параметров из микшированного звукового сигнала; формирование параметров повторного 15 микширования из слепых параметров и параметров микширования; и применение параметров к микшированному сигналу, чтобы усовершенствовать один или более речевых сигналов источников в соответствии с параметрами микширования.

В некоторых реализациях, способ включает в себя формирование 20 пользовательского интерфейса для приема входных данных, задающих параметры микширования; получение параметра микширования через пользовательский интерфейс; получение первого звукового сигнала, включающего в себя сигналы источников; получение дополнительной информации, по меньшей мере, некоторая часть которой представляет зависимость между первым звуковым сигналом и одним 25 или более сигналами источников; и повторное микширование одного или более сигналов источников с использованием дополнительной информации и параметра микширования, чтобы сформировать второй звуковой сигнал.

В некоторых реализациях, способ включает в себя получение первого 30 многоканального звукового сигнала, содержащего набор объектов; получение дополнительной информации, по меньшей мере, некоторая часть которой представляет зависимость между первым многоканальный звуковым сигналом и одним или более сигналами источников, представляющими подмножество объектов, которые должны повторно микшироваться; получение набора параметров 35 микширования; и формирование второго многоканального звукового сигнала с использованием дополнительной информации и набора параметров микширования.

В некоторых реализациях, способ включает в себя получение микшированного звукового сигнала; получение набора параметров микширования для повторного 40 микширования микшированного звукового сигнала; формирование параметров повторного микширования с использованием микшированного звукового сигнала и набора параметров микширования; и формируют повторно микшированный звуковой сигнал, применяя параметры повторного микширования к микшированному звуковому сигналу с использованием матрицы  $n$  на  $n$ .

45 Раскрыты другие реализации для усовершенствования звукового сигнала возможностью повторного микширования, в том числе реализации, направленные на системы, способы, устройства, машиночитаемые носители и пользовательские интерфейсы.

50 Описание чертежей

Фиг. 1А - структурная схема реализации системы кодирования для кодирования стереофонического сигнала плюс  $M$  сигналов источников, соответствующих объектам, которые должны повторно микшироваться в декодере.

Фиг.1В - блок-схема последовательности операций способа реализации последовательности операций для кодирования стереофонического сигнала плюс  $M$  сигналов источников, соответствующих объектам, которые должны повторно микшироваться в декодере.

5 Фиг.2 иллюстрирует времячастотное графическое представление для анализа и обработки стереофонического сигнала и  $M$  сигналов источников.

Фиг.3А - структурная схема реализации системы повторного микширования для оценки повторно микшированного стереофонического сигнала с использованием  
10 исходного стереофонического сигнала плюс дополнительной информации.

Фиг.3В - блок-схема последовательности операций способа реализации последовательности операций для оценки повторно микшированного стереофонического сигнала с использованием системы повторного микширования по  
15 фиг.3А.

Фиг.4 иллюстрирует индексы  $i$  коэффициентов краткосрочного преобразования Фурье (STFT), принадлежащих сегменту с индексом  $b$ .

Фиг.5 иллюстрирует группирование спектральных коэффициентов равномерного спектра STFT в имитационное неравномерное частотное разрешение слухового  
20 аппарата человека.

Фиг.6А - структурная схема реализации системы кодирования по фиг.1, объединенной с традиционным стереофоническим звуковым кодировщиком.

Фиг.6В - блок-схема последовательности операций способа реализации последовательности операций кодирования с использованием системы кодирования  
25 по фиг.1, объединенной с традиционным стереофоническим звуковым кодировщиком.

Фиг.7А - структурная схема реализации системы повторного микширования по фиг.3А, объединенной с традиционным стереофоническим звуковым декодером.

Фиг.7В - блок-схема последовательности операций способа реализации последовательности операций повторного микширования с использованием системы  
30 повторного микширования по фиг.7А, объединенной со стереофоническим звуковым кодировщиком.

Фиг.8А - структурная схема реализации системы кодирования, реализующей полностью слепое формирование дополнительной информации.

Фиг.8В - блок-схема последовательности операций способа реализаций последовательности операций кодирования с использованием системы кодирования  
35 по фиг.8А.

Фиг.9 иллюстрирует примерную функцию усиления,  $f(M)$ , для требуемого перепада  
40 уровней источников,  $L_i=L$  dB.

Фиг.10 - схема реализации последовательности операций формирования дополнительной информации с использованием технологии частично слепого формирования.

Фиг.11 - структурная схема реализации клиент-серверной архитектуры для поставки  
45 стереофонических сигналов и  $M$  сигналов источников и/или дополнительной информации в звуковые устройства с возможность повторного микширования.

Фиг.12 иллюстрирует реализацию пользовательского интерфейса для медиаплеера с  
50 возможностью повторного микширования.

Фиг.13 иллюстрирует реализацию системы декодирования, объединяющей декодирование пространственно кодированных звуковых объектов (SAOC) и декодирование повторного микширования.

Фиг.14А иллюстрирует обычную модель микширования для отдельного уровня

громкости диалога (SDV).

Фиг.14В иллюстрирует реализацию системы, объединяющей SDV и технологию повторного микширования.

5 Фиг.15 иллюстрирует реализацию рендерера эквивалентного микширования, показанного на фиг.14В.

Фиг.16 иллюстрирует реализацию системы распространения для технологии повторного микширования, описанной со ссылкой на фиг.1-15.

10 Фиг.17А иллюстрирует элементы различных реализации битового потока для предоставления информации повторного микширования.

Фиг.17В иллюстрирует реализацию интерфейса кодировщика повторного микширования для формирования битовых потоков, проиллюстрированных на фиг.17А.

15 Фиг.17С иллюстрирует реализацию интерфейса декодера повторного микширования для приема битовых потоков, сформированных кодировщиком, проиллюстрированным на фиг.17В.

Фиг.18 - структурная схема реализации системы, включающей в себя расширения для формирования добавочной дополнительной информации для определенных объектных сигналов, чтобы обеспечивать улучшенные эксплуатационные качества повторного микширования.

Фиг.19 - структурная схема реализации формирователя (рендерера) повторно микшированного сигнала, показанного на фиг.18.

Подробное описание

25 I. Повторное микширование стереофонических сигналов

Фиг.1А - структурная схема реализации системы 100 кодирования для кодирования стереофонических сигналов плюс М сигналов источников, соответствующих объектам, которые должны повторно микшироваться в декодере. В некоторых реализациях, система 100 кодирования обычно включает в себя матрицу 102 гребенки фильтров, формирователь 104 дополнительной информации и кодировщик 106.

*А. Исходный и требуемый повторно микшированный сигнал*

35 Два канала стереофонического звукового сигнала с временной дискретизацией обозначены  $\tilde{x}_1(n)$  и  $\tilde{x}_2(n)$ , где  $n$  - индекс времени. Предполагается, что стереофонический сигнал может быть представлен в виде:

$$\tilde{x}_1(n) = \sum_{i=1}^I a_i \tilde{s}_i(n), \quad (1)$$

$$40 \tilde{x}_2(n) = \sum_{i=1}^I b_i \tilde{s}_i(n),$$

где  $I$  - количество сигналов источников (например, инструментов), которые содержатся в стереофоническом сигнале (например, МРЗ), а  $\tilde{s}_i(n)$  - сигналы источников. Коэффициенты  $a_i$  и  $b_i$  определяют коэффициент усиления и амплитудное панорамирование для каждого сигнала источника. Предполагается, что все сигналы источников являются взаимно независимыми. Сигналы источников могут не все быть чистыми сигналами источников. Предпочтительнее некоторые из сигналов источников могут содержать реверберацию и/или другие сигнальные составляющие звуковых эффектов. В некоторых реализациях, задержки,  $d_i$ , могут привноситься в исходный звуковой сигнал микширования в [1] для содействия временному выравниванию с параметрами повторного микширования:

$$\tilde{x}_1(n) = \sum_{i=1}^I a_i \tilde{s}_i(n - d_i), \quad (1.1)$$

$$\tilde{x}_2(n) = \sum_{i=1}^I b_i \tilde{s}_i(n - d_i).$$

5 В некоторых реализациях, система 100 кодирования выдает или формирует информацию (в дальнейшем также указываемую ссылкой как «дополнительная информация») для модифицирования исходного стереофонического звукового сигнала (в дальнейшем также указываемого ссылкой как «стереофонический сигнал»),  
10 из условия, чтобы  $M$  сигналов источников «повторно микшировались» в стереофонический сигнал с разными коэффициентами усиления. Требуемый модифицированный стереофонический сигнал может быть представлен в виде:

$$\tilde{y}_1(n) = \sum_{i=1}^M c_i \tilde{s}_i(n) + \sum_{i=M+1}^I a_i \tilde{s}_i(n), \quad (2)$$

$$\tilde{y}_2(n) = \sum_{i=1}^M d_i \tilde{s}_i(n) + \sum_{i=M+1}^I b_i \tilde{s}_i(n),$$

15 где  $c_i$  и  $d_i$  - новые коэффициенты усиления (в дальнейшем также указываемые ссылкой как «коэффициенты усиления микширования» или «параметры микширования») для  $M$  сигналов источников, которые должны микшироваться (то есть сигналы источников с индексами 1, 2, ...,  $M$ ).

20 Цель системы 100 кодирования состоит в том, чтобы выдавать или формировать информацию для повторного микширования стереофонического сигнала, при заданном только исходном стереофоническом сигнале и небольшом количестве дополнительной информации (например, небольшой по сравнению с информацией, содержащейся в форме колебаний стереофонического сигнала).  
25 Дополнительная информация, поставляемая или формируемая системой 100 кодирования, может использоваться в декодере для относящейся к восприятию имитации требуемого модифицированного стереофонического сигнала [2] при заданном исходном стереофоническом сигнале [1].  
30 С системой 100 кодирования формирователь 104 дополнительной информации формирует дополнительную информацию для повторного микширования исходного стереофонического сигнала, а система 300 декодирования (фиг.3А) формирует требуемый повторно микшированный стереофонический звуковой сигнал с использованием дополнительной информации и исходного стереофонического сигнала.

#### 35 В. Обработка кодировщика

40 Вновь со ссылкой на фиг.1А исходный стереофонический сигнал и  $M$  сигналов источников подаются в качестве входного сигнала в матрицу 102 гребенки фильтров. Исходный стереофонический сигнал также непосредственно выводится из кодировщика 102. В некоторых реализациях, стереофонический сигнал,  
45 непосредственно выведенный из кодировщика 102, может задерживаться для синхронизации с битовым потоком дополнительной информации. В других реализациях, вывод стереофонического сигнала может синхронизироваться с дополнительной информацией в декодере. В некоторых реализациях, система 100 кодирования адаптируется к статистическим данным сигнала в качестве функции  
50 времени и частоты. Таким образом, для анализа и синтеза стереофонический сигнал и  $M$  сигналов источников обрабатываются во времячастотном представлении, как описано со ссылкой на фиг.4 и 5.

Фиг.1В - блок-схема последовательности операций способа реализации

последовательности 108 операций для кодирования стереофонического сигнала плюс  $M$  сигналов источников, соответствующих объектам, которые должны повторно микшироваться в декодере. Входной стереофонический сигнал и  $M$  сигналов источников раскладываются на поддиапазоны (110). В некоторых реализациях, разложение реализуется с помощью матрицы гребенки фильтров. Для каждого поддиапазона коэффициенты усиления оцениваются для  $M$  сигналов источников (112), как более полно описано ниже. Для каждого поддиапазона краткосрочные оценки мощности вычисляются для  $M$  сигналов источников (114), как описано ниже.

Оцененные коэффициенты усиления и мощности поддиапазонов могут квантоваться и кодироваться для формирования дополнительной информации (116).

Фиг.2 иллюстрирует времячастотное графическое представление для анализа и обработки стереофонического сигнала и  $M$  сигналов источников. Ось ординат графика представляет частоту и делится на множество неравномерных поддиапазонов 202. Ось абсцисс представляет время и делится на временные интервалы 204. Каждый из штрихпунктирных прямоугольников на фиг.2 представляет соответствующую пару поддиапазона и временного интервала. Таким образом, для заданного временного интервала 204 один или более поддиапазонов 202 соответствующих временному интервалу 204, могут обрабатываться в качестве группы 206. В некоторых реализациях, ширины поддиапазонов 202 выбираются на основании пределов восприятия, ассоциативно связанных со слуховым аппаратом человека, как описано со ссылкой на фиг.4 и 5.

В некоторых реализациях, входной стереофонический сигнал и  $M$  входных сигналов источников раскладываются матрицей 102 гребенки фильтров на некоторое количество поддиапазонов 202. Поддиапазоны 202 на каждой центральной частоте могут обрабатываться подобным образом. Пара поддиапазонов стереофонических звуковых входных сигналов, на отдельной частоте, обозначена  $x_1(k)$  и  $x_2(k)$ , где  $k$  подвергнутый понижающей дискретизации индекс времени сигналов поддиапазонов. Подобным образом соответствующие сигналы поддиапазонов  $M$  входных сигналов источников обозначены  $s_1(k), s_2(k), \dots, s_M(k)$ . Отметим, что для простоты обозначения, индексы для поддиапазонов были опущены в этом примере. Что касается понижающей дискретизации, для эффективности могут использоваться сигналы поддиапазонов с меньшей частотой дискретизации. Обычно гребенки фильтров и STFT фактически получают подвергнутые подвыборке сигналы (или спектральные коэффициенты).

В некоторых реализациях, дополнительная информация, необходимая для повторного микширования сигнала источника с индексом  $i$ , включает в себя коэффициенты  $a_i$  и  $b_i$  усиления и в каждом поддиапазоне, оценку мощности сигнала поддиапазона в качестве функции времени,  $E\{s_i^2(k)\}$ . Коэффициенты  $a_i$  и  $b_i$  усиления могут задаваться (если эти сведения о стереофоническом сигнале известны) или оцениваться. Для многих стереофонических сигналов,  $a_i$  и  $b_i$  являются статическими. Если  $a_i$  и  $b_i$  являются меняющимися в качестве функции времени  $k$ , эти коэффициенты усиления могут оцениваться в качестве функции времени. Она не обязательно должна использовать среднее значение или оценку мощности поддиапазона для формирования дополнительной информации. Предпочтительнее в некоторых реализациях реальная мощность  $S_i^2$  поддиапазона может использоваться в качестве оценки мощности.

В некоторых реализациях, краткосрочная мощность поддиапазона может

оцениваться с использованием однополюсного усреднения, где  $E\{s_i^2(k)\}$  может вычисляться как:

$$E\{s_i^2(k)\} = \alpha s_i^2(k) + (1 - \alpha)E\{s_i^2(k-1)\}, \quad (3)$$

5 где  $\alpha \in [0,1]$  определяет постоянную времени экспоненциально затухающего окна оценки,

$$T = \frac{1}{\alpha f_s}, \quad (4)$$

а  $f_s$  обозначает частоту дискретизации поддиапазона. Подходящим значением для  $T$ ,  
10 например, может быть 40 миллисекунд. В последующих уравнениях  $E\{\cdot\}$  обычно обозначает краткосрочное усреднение.

В некоторых реализациях, некоторая или вся из дополнительной информации  $a_i$ ,  $b_i$  и  $E\{s_i^2(k)\}$  может поставляться на тех же носителях, что и стереофонический сигнал.  
15 Например, издатель музыкальных произведений, студия звукозаписи, ретушер звукозаписи или тому подобные, могут поставлять дополнительную информацию с соответствующим стереофоническим сигналом на компакт-диске (CD), цифровом видеодиске (DVD), флэш-памяти и тому подобном. В некоторых реализациях, некоторая или вся из дополнительной информации может поставляться по сети  
20 (например, сети Интернет, сети Ethernet, беспроводной сети) встраиванием дополнительной информации в битовый поток стереофонического сигнала или передачей дополнительной информации в отдельном битовом потоке.

Если  $a_i$  и  $b_i$  не заданы, то эти коэффициенты могут оцениваться. Поскольку,  
25  $E\{\tilde{s}_i(n)\tilde{x}_1(n)\} = a_i E\{\tilde{s}_i^2(n)\}$ ,  $a_i$  может вычисляться как:

$$a_i = \frac{E\{\tilde{s}_i(n)\tilde{x}_1(n)\}}{E\{\tilde{s}_i^2(n)\}}. \quad (5)$$

Подобным образом  $b_i$  может вычисляться в качестве

$$30 \quad b_i = \frac{E\{\tilde{s}_i(n)\tilde{x}_2(n)\}}{E\{\tilde{s}_i^2(n)\}}. \quad (6)$$

Если  $a_i$  и  $b_i$  являются самонастраивающимися со временем, оператор  $E\{\cdot\}$  представляет операцию краткосрочного усреднения. С другой стороны, если  
35 коэффициенты  $a_i$  и  $b_i$  усиления являются статическими, коэффициенты усиления могут вычисляться, учитывая стереофонические сигналы во всей их полноте. В некоторых вариантах осуществления коэффициенты  $a_i$  и  $b_i$  усиления могут оцениваться независимо для каждого поддиапазона. Отметим, что в [5] и [6] сигналы  $s_i$  источников  
40 являются независимыми, но, вообще, не сигнал  $s_i$  источника и стереофонические каналы  $x_1$  и  $x_2$ , поскольку  $s_i$  содержится в стереофонических каналах  $x_1$  и  $x_2$ .

В некоторых реализациях краткосрочные оценки мощности и коэффициенты усиления для каждого поддиапазона квантуются и кодируются кодировщиком 106,  
45 чтобы сформировать дополнительную информацию (например, битовый поток с низкой скоростью передачи битов). Отметим, что эти значения могут не квантоваться и кодироваться непосредственно, но сначала могут преобразовываться в другие величины, более подходящие для квантования и кодирования, как описано со ссылкой на фиг.4 и 5. В некоторых реализациях,  $E\{s_i^2(k)\}$  может быть нормализована  
50 относительно мощности поддиапазона входного стереофонического звукового сигнала, делая систему 100 кодирования устойчивой касательно изменений, когда традиционный звуковой кодировщик используется для эффективного кодирования стереофонического звукового сигнала, как описано со ссылкой на фиг.6-7.

### *С. Обработка декодера*

Фиг.3А - структурная схема реализации системы 300 повторного микширования для оценки повторно микшированного стереофонического сигнала с использованием исходного стереофонического сигнала плюс дополнительной информации. В некоторых реализациях, система 300 повторного микширования обычно включает в себя матрицу 302 гребенки фильтров, декодер 304, модуль 306 повторного микширования и обратную матрицу 308 гребенки фильтров.

Оценка повторно микшированного звукового сигнала может выполняться независимо в некотором количестве поддиапазонов. Дополнительная информация включает в себя мощность поддиапазона,  $E\{s_i^2(k)\}$  и коэффициенты усиления,  $a_i$  и  $b_i$ , с которыми  $M$  сигналов источников содержатся в стереофоническом сигнале. Новые коэффициенты усиления или коэффициенты усиления микширования *требуемого* повторно микшированного стереофонического сигнала представлены посредством  $c_i$  и  $d_i$ . Коэффициенты усиления  $c_i$  и  $d_i$  микширования могут задаваться пользователем через пользовательский интерфейс звукового устройства, такого как описанное со ссылкой на фиг.12.

В некоторых реализациях входной стереофонический сигнал раскладывается на поддиапазоны матрицей 302 гребенки фильтров, где пара поддиапазонов на отдельной частоте обозначена  $x_1(k)$  и  $x_2(k)$ . Как проиллюстрировано на фиг.3А, дополнительная информация декодируется декодером 304, давая для каждого из  $M$  сигналов источников, которые должны повторно микшироваться, коэффициенты  $a_i$  и  $b_i$  усиления, которые содержатся во входном стереофоническом сигнале, и для каждого поддиапазона оценку мощности,  $E\{s_i^2(k)\}$ . Декодирование дополнительной информации описано более подробно со ссылкой на фиг.4 и 5.

При заданной дополнительной информации, соответствующая пара поддиапазонов повторно микшированного стереофонического звукового сигнала может оцениваться модулем 306 повторного микширования в качестве функции коэффициентов  $c_i$  и  $d_i$  усиления микширования повторно микшированного стереофонического сигнала. Обратная матрица 308 гребенки фильтров применяется к оцененным парам поддиапазонов для предоставления повторно микшированного стереофонического сигнала временной области.

Фиг.3В - блок-схема последовательности операций способа реализации последовательности 310 операций для оценивания повторно микшированного стереофонического сигнала с использованием системы повторного микширования по фиг.3А. Входной стереофонический сигнал раскладывается на пары поддиапазонов (312). Дополнительная информация декодируется для пар поддиапазонов (314). Пары поддиапазонов повторно микшируются с использованием дополнительной информации и коэффициентов усиления микширования (318). В некоторых реализациях, коэффициенты усиления микширования предоставляются пользователем, как описано со ссылкой на фиг.12. В качестве альтернативы коэффициенты усиления микширования могут предоставляться программно, приложением, операционной системой или тому подобным. Коэффициенты усиления микширования также могут поставляться по сети (например, сети Интернет, сети Ethernet, беспроводной сети), как описано со ссылкой на фиг.11.

### *Д. Последовательность операций повторного микширования*

В некоторых реализациях, повторно микшированный стереофонический сигнал может аппроксимироваться в математическом смысле с использованием оценки методом наименьших квадратов. По выбору относящиеся к восприятию соображения

могут использоваться для модификации оценки.

Уравнения [1] и [2] к тому же рассматривают пары  $x_1(k)$  и  $x_2(k)$ , а также  $y_1(k)$  и  $y_2(k)$  поддиапазонов соответственно. В этом случае сигналы источников замещаются сигналами поддиапазонов источников,  $s_i(k)$ .

5 Пара поддиапазонов стереофонического сигнала задается согласно

$$x_1(k) = \sum_{i=1}^I a_i s_i(k) \quad (7)$$

$$10 \quad x_2(k) = \sum_{i=1}^I b_i s_i(k)$$

а парой поддиапазонов повторно микшированного стереофонического звукового сигнала являются

$$15 \quad y_1(k) = \sum_{i=1}^M c_i s_i(k) + \sum_{i=M+1}^I a_i s_i(k), \quad (8)$$

$$y_2(k) = \sum_{i=1}^M d_i s_i(k) + \sum_{i=M+1}^I b_i s_i(k)$$

20 При заданной паре поддиапазона исходного стереофонического сигнала,  $x_1(k)$  и  $x_2(k)$ , пара поддиапазона стереофонического сигнала с разными коэффициентами усиления оценивается в качестве линейной комбинации исходной пары левого и правого стереофонических поддиапазонов:

$$\tilde{y}_1(k) = w_{11}(k)x_1(k) + w_{12}(k)x_2(k), \quad (9)$$

$$25 \quad \tilde{y}_2(k) = w_{21}(k)x_1(k) + w_{22}(k)x_2(k),$$

где  $w_{11}(k)$ ,  $w_{12}(k)$ ,  $w_{21}(k)$  и  $w_{22}(k)$  - заданные действительными значениями весовые коэффициенты.

30 Ошибка оценки определяется в качестве

$$e_1(k) = y_1(k) - \hat{y}_1(k),$$

$$= y_1(k) - w_{11}(k)x_1(k) - w_{12}(k)x_2(k),$$

(10)

$$35 \quad = y_2(k) - w_{21}(k)x_1(k) - w_{22}(k)x_2(k),$$

$$e_2(k) = y_2(k) - \hat{y}_2(k).$$

40 Веса  $w_{11}(k)$ ,  $w_{12}(k)$ ,  $w_{21}(k)$  и  $w_{22}(k)$  могут вычисляться в каждый момент  $k$  времени для поддиапазонов на каждой частоте, из условия чтобы минимизировались среднеквадратические ошибки  $E\{e_1^2(k)\}$  и  $E\{e_2^2(k)\}$ . Для вычисления  $w_{11}(k)$  и  $w_{12}(k)$ , отметим, что  $E\{e_1^2(k)\}$  минимизируется, когда ошибка  $e_1(k)$  ортогональна  $x_1(k)$  и  $x_2(k)$ , то есть

$$E\{(y_1 - w_{11}x_1 - w_{12}x_2)x_1\} = 0$$

$$45 \quad E\{(y_1 - w_{11}x_1 - w_{12}x_2)x_2\} = 0.$$

(11)

Отметим, что для удобства обозначения индекс  $k$  времени был опущен.

Переписывание этих уравнений дает:

$$E\{x_1^2\}w_{11} + E\{x_1x_2\}w_{12} = E\{x_1y_1\},$$

(12)

$$50 \quad E\{x_1x_2\}w_{11} + E\{x_2^2\}w_{12} = E\{x_2y_1\}.$$

Коэффициенты усиления являются решением этой системы линейных уравнений:

$$w_{11} = \frac{E\{x_2^2\}E\{x_1y_1\} - E\{x_1x_2\}E\{x_2y_1\}}{E\{x_1^2\}E\{x_2^2\} - E^2\{x_1x_2\}}, \quad (13)$$

$$w_{12} = \frac{E\{x_1x_2\}E\{x_1y_1\} - E\{x_1^2\}E\{x_2y_1\}}{E^2\{x_1x_2\} - E\{x_1^2\}E\{x_2^2\}}.$$

5

Несмотря на то, что  $E\{x_1^2\}$ ,  $E\{x_2^2\}$  и  $E\{x_1x_2\}$  могут непосредственно оцениваться при заданной паре поддиапазонов входного стереофонического сигнала декодера,  $E\{x_1y_1\}$  и  $E\{x_2y_2\}$  могут оцениваться с использованием дополнительной информации ( $E\{s_1^2\}$ ,  $a_1$ ,  $b_1$ ) и коэффициентов усиления микширования,  $c_1$  и  $d_1$ , требуемого повторно микшированного стереофонического сигнала:

$$E\{x_1y_1\} = E\{x_1^2\} + \sum_{i=1}^M a_i(c_i - a_i)E\{s_i^2\}, \quad (14)$$

15

$$E\{x_2y_1\} = E\{x_1x_2\} + \sum_{i=1}^M b_i(c_i - a_i)E\{s_i^2\}.$$

Подобным образом вычисляются  $w_{21}$  и  $w_{22}$ , давая в результате

$$w_{21} = \frac{E\{x_2^2\}E\{x_1y_2\} - E\{x_1x_2\}E\{x_2y_2\}}{E\{x_1^2\}E\{x_2^2\} - E^2\{x_1x_2\}}, \quad (15)$$

20

$$w_{22} = \frac{E\{x_1x_2\}E\{x_1y_2\} - E\{x_1^2\}E\{x_2y_2\}}{E^2\{x_1x_2\}E\{x_2^2\} - E\{x_1^2\}E\{x_2^2\}}.$$

Причем

25

$$E\{x_2y_2\} = E\{x_2^2\} + \sum_{i=1}^M b_i(d_i - b_i)E\{s_i^2\}. \quad (16)$$

Когда левый и правый сигналы поддиапазонов когерентны или почти когерентны, то есть когда

30

$$\phi = \frac{E\{x_1x_2\}}{\sqrt{E\{x_1^2\}E\{x_2^2\}}} \quad (17)$$

близко к единице, то решение для весов неуникально или плохо обусловлено. Таким образом, если  $\phi$  является большим, чем определенное пороговое значение (например, 0,95), то веса, например, вычисляются согласно

35

$$w_{11} = \frac{E\{x_1y_1\}}{E\{x_1^2\}},$$

$$w_{12} = w_{21} = 0,$$

(18)

40

$$w_{22} = \frac{E\{x_2y_2\}}{E\{x_2^2\}}.$$

При условии  $\phi=1$  уравнение [18] является одним из неуникальных решений, удовлетворяющих [12] и подобной системе уравнений ортогональности для двух других весов. Отметим, что когерентность в [17] используется для вынесения суждения, насколько  $x_1$  и  $x_2$  подобны друг другу. Если когерентность равна нулю, то  $x_1$  и  $x_2$  независимы. Если когерентность равна единице, то  $x_1$  и  $x_2$  подобны (но могут иметь разные уровни). Если  $x_1$  и  $x_2$  сильно подобны (когерентность близка к единице), то двухканальное вычисление Винера (вычисление четырех весов) является плохо обусловленным. Примерным диапазоном для порогового значения является от приблизительно 0,4 до приблизительно 1,0.

50

Результирующий повторно микшированный стереофонический сигнал, полученный преобразованием вычисленных сигналов поддиапазонов во временную область,

звучание, подобное стереофоническому сигналу, который действительно микшировался бы с разными коэффициентами усиления микширования,  $c_i$  и  $d_i$ , (в последующем этот сигнал обозначен «требуемым сигналом»). С одной стороны, математически это требует, чтобы вычисленные сигналы поддиапазонов были подобны действительно по-разному микшированным сигналам поддиапазонов. Это верно до определенной степени. Поскольку оценка выполняется в области мотивированных восприятием поддиапазонов, требование к подобию является менее строгим. До тех пор, пока контрольные сигналы значимого для восприятия определения местоположения (например, контрольные сигналы перепада уровней и когерентности) достаточно подобны, вычисленный повторно микшированный стереофонический сигнал будет звучать подобно требуемому сигналу.

*Е. Необязательная возможность: настройка контрольных сигналов перепадов уровней*

В некоторых реализациях, если используется обработка, описанная в материалах настоящей заявки, могут быть получены хорошие результаты. Тем не менее, чтобы быть уверенным, что важные контрольные сигналы определения местоположения перепада уровней весьма близки к контрольным сигналам перепадов уровней требуемого сигнала, домасштабирование поддиапазонов может применяться для «настройки» контрольных сигналов перепадов уровней, чтобы удостовериться, что они соответствуют контрольным сигналам перепадов уровней требуемого сигнала.

Для модификации оценок сигналов поддиапазонов методом наименьших квадратов в [9] учитывается мощность поддиапазона. Если мощность поддиапазона является надлежащей, то перепад уровней важного пространственного контрольного сигнала также может быть надлежащим. Требуемой мощностью левого поддиапазона сигнала [8] является

$$E\{y_1^2\} = E\{x_1^2\} + \sum_{i=1}^M (c_i^2 - a_i^2) E\{s_i^2\}, \quad (19)$$

а мощностью поддиапазона оценки по [9] является

$$E\{\hat{y}_1^2\} = E\{(w_{11}x_1 + w_{12}x_2)^2\} \\ = w_{11}^2 E\{x_1^2\} + 2w_{11}w_{12} E\{x_1x_2\} + w_{12}^2 E\{x_2^2\}. \quad (20)$$

Таким образом, чтобы  $\hat{y}_1(k)$  имел такую же мощность, как  $y_1(k)$ , он должен перемножаться с

$$g_1 = \sqrt{\frac{E\{x_1^2\} + \sum_{i=1}^M (c_i^2 - a_i^2) E\{s_i^2\}}{w_{11}^2 E\{x_1^2\} + 2w_{11}w_{12} E\{x_1x_2\} + w_{12}^2 E\{x_2^2\}}}. \quad (21)$$

Подобным образом  $\hat{y}_2(k)$  перемножается с

$$g_2 = \sqrt{\frac{E\{x_2^2\} + \sum_{i=1}^M (d_i^2 - b_i^2) E\{s_i^2\}}{w_{21}^2 E\{x_1^2\} + 2w_{21}w_{22} E\{x_1x_2\} + w_{22}^2 E\{x_2^2\}}}, \quad (22)$$

чтобы иметь такую же мощность, как требуемый сигнал  $y_2(k)$  поддиапазона.

## II. Квантование и кодирование дополнительной информации

### A. Кодирование

Как описано в предыдущем разделе, дополнительной информацией, необходимой для повторного микширования сигнала источника с индексом  $i$ , являются коэффициенты  $a_i$  и  $b_i$ , и в каждом поддиапазоне мощность как функция времени,  $E\{s_i^2(k)\}$ . В некоторых реализациях, соответствующие значения перепадов усиления и

уровней для коэффициентов  $a_i$  и  $b_i$  усиления могут вычисляться в дБ, как изложено ниже:

$$g_i = 10 \log_{10} (a_i^2 + b_i^2),$$

5

(23)

$$l_i = 20 \log_{10} \frac{b_i}{a_i}.$$

10

В некоторых реализациях, значения усиления и перепада уровней квантуются и кодируются кодом Хаффмана. Например, равномерный квантователь с размером шага квантователя в 2 дБ и одномерный кодировщик Хаффмана могут использоваться для квантования и кодирования соответственно. Другие известные квантователи и кодировщики также могут использоваться (например, векторный квантователь).

15

Если  $a_i$  и  $b_i$  являются независимыми от времени, и предполагается, что дополнительная информация достоверно поступает в декодер, соответствующим кодированным значениям необходимо передаваться только один раз. Иначе  $a_i$  и  $b_i$  могут передаваться через равные промежутки времени или в ответ на запускающее событие (например, всякий раз, когда меняются кодированные значения).

20

Чтобы быть устойчивым к масштабированию стереофонического сигнала и потере/приросту мощности, обусловленным кодированием стереофонического сигнала, в некоторых реализациях мощность  $E\{s_i^2(k)\}$  поддиапазона не кодируется непосредственно в качестве дополнительной информации. Предпочтительнее может использоваться показатель, определенный относительно стереофонического сигнала:

25

$$A_i(k) = 10 \log_{10} \frac{E\{s_i^2(k)\}}{E\{x_1^2(k)\} + E\{x_2^2(k)\}}. \quad (24)$$

30

Может быть полезным использовать одинаковые окна/постоянные времени оценки для вычисления  $E\{\cdot\}$  по разным сигналам. Преимущество определения дополнительной информации в качестве относительного значения мощности [24] содержит этапы, на которых в декодере, если требуется, могут использоваться окно/постоянная времени оценки иные, чем в кодировщике. К тому же влияние временной рассогласованности между дополнительной информацией и стереофоническим сигналом снижается по сравнению со случаем, когда мощность источника передавалась бы в качестве абсолютного значения. Для квантования и кодирования  $A_i(k)$  в некоторых реализациях используется равномерный квантователь, например, с размером шага 2 дБ, и одномерный кодировщик Хаффмана. Результирующая скорость передачи битов может быть такой же небольшой, как 3 килобита/с (килобитов в секунду) на звуковой объект, который должен повторно микшироваться.

35

40

В некоторых реализациях, скорость передачи битов может снижаться, когда входной сигнал источника, соответствующий объекту, который должен повторно микшироваться в декодере, является неозвученным. Режим кодирования кодировщика может обнаруживать неозвученный объект, а затем передавать в декодер информацию (например, единственный бит на кадр) для указания, что объект является неозвученным.

45

### *В. Декодирование*

50

При заданных декодированных из кода Хаффмана (квантованных) значениях [23] и [24] значения, необходимые для повторного микширования, могут вычисляться, как изложено ниже:

$$\tilde{a}_i = \frac{10^{\frac{\tilde{a}_i}{20}}}{\sqrt{1 + 10^{\frac{i}{10}}}},$$

5

$$\tilde{b}_i = \frac{10^{\frac{\tilde{a}_i + i}{20}}}{\sqrt{1 + 10^{\frac{i}{10}}}}, \quad (25)$$

$$\hat{E}\{s_i^2(k)\} = 10^{\frac{\tilde{A}_i(k)}{10}} (E\{x_1^2(k)\} + E\{x_2^2(k)\}).$$

10

III. Подробности реализации

#### A. Времячастотная обработка

15

В некоторых реализациях, основанная на STFT (краткосрочном преобразовании Фурье) обработка используется для систем кодирования/декодирования, описанных со ссылкой на фиг.1-3. Могут использоваться другие времячастотные преобразования для достижения требуемого результата, в том числе, но не в качестве ограничения, гребенка фильтров с квадратурными зеркальными фильтрами (QMF), модифицированное дискретное косинусное преобразование (MDCT), гребенка вейвлетных фильтров и т.п.

20

Для обработки анализа (например, прямой операции гребенки фильтров) в некоторых реализациях кадр из  $N$  отсчетов может перемножаться с окном до того, как применяется  $N$ -точечное дискретное преобразование Фурье (ДПФ, DFT) или быстрое преобразование Фурье (БПФ, FFT). В некоторых реализациях, может использоваться следующее синусоидальное окно:

25

$$w_a(l) = \begin{cases} \sin\left(\frac{n\pi}{N}\right) & \text{для } 0 \leq n < N \\ 0 & \text{иначе.} \end{cases} \quad (26)$$

30

Если размер блока обработки является иным, чем размер ДПФ/БПФ, то в некоторых реализациях заполнение нулями может использоваться для эффективного получения меньшего окна, чем  $N$ . Описанная обработка анализа, например, может повторяться каждые  $N/2$  отсчетов (размер скачка окна с равными количествами), давая в результате 50-процентное перекрытие окон. Другие вырезающие (оконные) функции и процентные перекрытия могут использоваться для достижения требуемого

35

результата. Для преобразования из спектральной области STFT во временную область к спектрам могут применяться обратные ДПФ или БПФ. Результирующий сигнал вновь перемножается с окном, описанным в [26], и смежные блоки сигналов, являющиеся следствием перемножения с окном, объединяются с перекрытием, добавленным для получения непрерывного сигнала во временной области.

40

В некоторых случаях равномерное спектральное разрешение STFT может не быть хорошо адаптированным к человеческому восприятию. В таких случаях в противоположность обработке каждого частотного коэффициента STFT по отдельности коэффициенты STFT могут «группироваться» из условия, чтобы одна группа имела ширину полосы пропускания, приблизительно в два раза большую эквивалентной прямоугольной полосы пропускания (ERB), каковое является пригодным частотным разрешением для пространственной звуковой обработки.

45

Фиг.4 иллюстрирует индексы  $i$  коэффициентов STFT, принадлежащих сегменту с индексом  $b$ . В некоторых реализациях, учитываются только первые  $N/2+1$  спектральных коэффициентов спектра, так как спектр является симметричным. Индексами коэффициентов STFT, которые принадлежат сегменту с индексом  $b$  ( $1 \leq b \leq B$ ), являются  $i \in \{A_{b-1}, A_{b-1}+1, \dots, A_b\}$ , причем  $A_0=0$ , как проиллюстрировано на

50

фиг.4. Сигналы, представленные спектральными коэффициентами сегментов, соответствуют мотивированному восприятием разложению поддиапазонов, используемому системой кодирования. Таким образом, в пределах каждого такого сегмента описанная обработка применяется совместно к коэффициентам STFT в пределах сегмента.

Фиг.5 примерно иллюстрирует группирование спектральных коэффициентов равномерного спектра STFT в имитационное неравномерное частотное разрешение слухового аппарата человека. На фиг.5  $N=1024$  для частоты дискретизации в 44,1 кГц и количества сегментов,  $B = 20$ , причем каждый сегмент имеет ширину полосы пропускания приблизительно в 2 ERB. Отметим, что последний сегмент является меньшим, чем ERB, вследствие среза на частоте Найквиста.

#### *В. Оценка статистических данных*

При заданных двух коэффициентах STFT,  $x_i(k)$  и  $x_j(k)$ , значения  $E\{x_i(k)x_j(k)\}$ , необходимые для вычисления повторно микшированного стереофонического звукового сигнала, могут оцениваться итерационно. В этом случае частота  $f_s$  дискретизации поддиапазона является частотой во времени, с которой вычисляются спектры STFT. Чтобы получать оценки для каждого относящегося к восприятию сегмента (а не для каждого коэффициента STFT), оцененные значения могут усредняться в пределах сегментов перед дальнейшим использованием.

Обработка, описанная в предыдущих разделах, может применяться к каждому сегменту, как будто он был одним диапазоном. Сглаживание между сегментами, например, может достигаться с использованием перекрывающихся спектральных окон, чтобы избегать внезапных изменений обработки по частоте, таким образом уменьшая артефакты.

#### *С. Сочетание с традиционными звуковыми кодировщиками*

Фиг.6А - структурная схема реализации системы 100 кодирования по фиг.1А, объединенной с традиционным стереофоническим звуковым кодировщиком. В некоторых реализациях комбинированная система 600 кодирования включает в себя традиционный звуковой кодировщик 602, предложенный кодировщик 604 (например, систему 100 кодирования) и объединитель 606 битовых потоков. В показанном примере стереофонические звуковые входные сигналы кодируются традиционным звуковым кодировщиком 602 (например, MP3, AAC, объемного звучания MPEG, и т.д.) и анализируются предложенным кодировщиком 604 для предоставления дополнительной информации, как описано ранее со ссылкой на фиг.1-5. Два результирующих битовых потока объединяются объединителем 606 потоков для предоставления обратно совместимого битового потока. В некоторых реализациях объединение результирующих битовых потоков включает в себя встраивание дополнительной информации с низкой скоростью передачи битов (например, коэффициентов  $a_i$ ,  $b_i$  усиления, и мощности  $E\{s_i^2(k)\}$  поддиапазона) в обратно совместимый битовый поток.

Фиг.6В - блок-схема последовательности операций способа реализации последовательности 608 операций кодирования с использованием системы 100 кодирования по фиг.1, объединенной с традиционным стереофоническим звуковым кодировщиком. Входной стереофонический сигнал кодируется с использованием традиционного стереофонического звукового кодировщика (610). Дополнительная информация формируется из стереофонического сигнала и  $M$  сигналов источников с использованием системы 100 кодирования по фиг.1А (612). Формируются (614) один или более обратно совместимых битовых потоков, включающих в себя кодированный

стереофонический сигнал и дополнительную информацию.

Фиг.7А - структурная схема реализации системы 300 повторного микширования по фиг.3А, объединенной с традиционным стереофоническим звуковым декодером для предоставления комбинированной системы 700. В некоторых реализациях комбинированная система 700 обычно включает в себя анализатор 702 битового потока, традиционный звуковой декодер 704 (например, МРЗ, ААС) и предложенный декодер 706. В некоторых реализациях предложенным декодером 706 является система 300 повторного микширования по фиг.3А.

В показанном примере битовый поток разделяется на стереофонический звуковой битовый поток и битовый поток, содержащий дополнительную информацию, необходимую предложенному декодеру 706 для обеспечения возможности повторного микширования. Стереофонический сигнал декодируется традиционным звуковым декодером 704 и подается в предложенный декодер 706, который модифицирует стереофонический сигнал в качестве функции дополнительной информации, полученной из битового потока и пользовательского ввода (например, коэффициентов  $c_i$  и  $d_i$  усиления микширования).

Фиг.7В - блок-схема последовательности операций способа реализации последовательности 708 операций повторного микширования с использованием комбинированной системы 700 по фиг.7А. Битовый поток, принятый из кодировщика, анализируется для выдачи битового потока кодированного стереоскопического сигнала и битового потока дополнительной информации (710). Кодированный стереофонический сигнал декодируется с использованием традиционного звукового декодера (712). Примерные декодеры включают в себя МРЗ, ААС (в том числе различные стандартизованные профили ААС), параметрический стереофонический сигнал, спектральную репликацию полос (SBR), объемное звучание MPEG или любую их комбинацию. Декодированный стереофонический сигнал повторно микшируется с использованием дополнительной информации и пользовательского ввода (например,  $c_i$  и  $d_i$ ).

#### IV. Повторное микширование многоканальных звуковых сигналов

В некоторых реализациях системы 100, 300 кодирования и повторного микширования, описанные в предыдущих разделах, могут быть расширены для повторного микширования многоканальных звуковых сигналов (например, сигналов объемного звучания 5,1). В дальнейшем стереофонический сигнал и многоканальный сигнал также указываются ссылкой как «многоканальные» сигналы. Специалисты в данной области техники поняли бы, каким образом следует переписать с [7] по [22] для схемы многоканального кодирования/декодирования, то есть для более чем двух сигналов  $x_1(k)$ ,  $x_2(k)$ ,  $x_3(k)$ , ...,  $x_C(k)$ , где  $C$  - количество звуковых каналов микшированного сигнала.

Уравнение [9] для многоканального случая становится:

$$\begin{aligned} \hat{y}_1(k) &= \sum_{c=1}^C w_{1c}(k)x_c(k), \\ \hat{y}_2(k) &= \sum_{c=1}^C w_{2c}(k)x_c(k), \\ &\dots \\ \hat{y}_C(k) &= \sum_{c=1}^C w_{Cc}(k)x_c(k). \end{aligned} \quad (27)$$

Уравнение, подобное [11] с  $C$  уравнениями, может выводиться и решаться для определения весов, как описано ранее.

В некоторых реализациях определенные каналы могут быть оставлены

необработанными. Например, для объемного звучания 5,1 два задних канала могут быть оставлены необработанными, а повторное микширование применяется только к передним левому, правому и центральному каналам. В этом случае алгоритм повторного микширования трех каналов может применяться к передним каналам.

5 Качество звукозаписи, вытекающее из раскрытой схемы повторного микширования, зависит от сущности модификации, которая выполняется. Для относительно слабых модификаций, например изменения панорамирования с 0 дБ до 15 дБ или модификации усиления в 10 дБ, результирующее качество звукозаписи 10 может быть более высоким, чем достигаемое традиционными технологиями. К тому же качество предложенной раскрытой схемы повторного микширования может быть выше, чем традиционные схемы повторного микширования, так как стереофонический сигнал модифицируется только по необходимости, чтобы достигать требуемого повторного микширования.

15 Схема повторного микширования, раскрытая в материалах настоящей заявки, предоставляет несколько преимуществ над традиционными технологиями. Прежде всего она предоставляет возможность повторного микширования меньшего, чем суммарное количества объектов в данном стереофоническом или многоканальном звуковом сигнале. Это достигается посредством оценивания дополнительной 20 информации в качестве функции данного стереофонического звукового сигнала, плюс  $M$  сигналов источников, представляющих  $M$  объектов в стереофоническом звуковом сигнале, которые должны быть задействованы для повторного микширования в декодере. Раскрытая система повторного микширования обрабатывает заданный 25 стереофонический сигнал в качестве функции дополнительной информации и в качестве функции пользовательского ввода (требуемого повторного микширования), чтобы формировать стереофонический сигнал, который подобен по восприятию стереофоническому сигналу, микшированному действительно по-иному.

30 V. Усовершенствования к основной схеме повторного микширования

#### *A. Предварительная обработка дополнительной информации*

Когда поддиапазон ослабляется слишком сильно относительно соседних поддиапазонов, могут возникать звуковые артефакты. Таким образом, требуется 35 ограничивать максимальное затухание. Более того, поскольку статистические данные стереофонического сигнала и объектного сигнала источника измеряются независимо в кодировщике и декодере соответственно, отношение между измеренной мощностью поддиапазона стереофонического сигнала и мощностью поддиапазона объектного сигнала (которое представлено дополнительной информацией) может отклоняться от 40 реальности. Вследствие этого дополнительная информация может быть такой, что она физически невозможна, например мощность сигнала у повторно микшированного сигнала [19] может становиться отрицательной. Обе из этих проблем могут быть преодолены, как описано ниже.

45 Мощностью поддиапазона левого и правого повторно смешанного сигнала является:

$$E\{y_1^2\} = E\{x_1^2\} + \sum_{i=1}^M (c_i^2 - a_i^2) P_{s_i},$$

$$E\{y_2^2\} = E\{x_2^2\} + \sum_{i=1}^M (d_i^2 - b_i^2) P_{s_i},$$
(28)

50 где  $P_{S_i}$  равна квантованной и закодированной оценке мощности поддиапазона, заданной в [25], которая вычисляется в качестве функции дополнительной информации. Мощность поддиапазона повторно микшированного сигнала может

быть ограничена так, что она никогда не бывает меньше, чем на  $L$  дБ ниже мощности поддиапазона исходного стереофонического сигнала,  $E\{x_1^2\}$ . Подобным образом  $E\{y_2^2\}$  ограничена, чтобы не быть меньше, чем на  $L$  дБ ниже  $E\{x_2^2\}$ . Этот результат может достигаться с помощью следующих операций:

1. Вычислить мощность поддиапазона левого и правого повторно микшированного сигнала согласно [28].

2. Если  $E\{y_1^2\} < QE\{x_1^2\}$ , то настроить вычисленные значения  $P_{Si}$  дополнительной информации из условия, чтобы сохранялось  $E\{y_1^2\} = QE\{x_1^2\}$ . Для ограничения мощности  $E\{y_1^2\}$ , чтобы никогда не была меньше, чем на  $A$  дБ ниже мощности  $E\{x_1^2\}$ ,  $Q$  может быть установлено в  $Q=10^{-A/10}$ . В таком случае  $P_{Si}$  может корректироваться ее перемножением с

$$\frac{(1-Q)E\{x_1^2\}}{-\sum_{i=1}^M (c_i^2 - a_i^2)P_i} \quad (29)$$

3. Если  $E\{y_2^2\} < QE\{x_2^2\}$ , то настроить вычисленные значения  $P_{Si}$  дополнительной информации из условия, чтобы сохранялось  $E\{y_2^2\} = QE\{x_2^2\}$ . Это может достигаться перемножением  $P_{Si}$  с

$$\frac{(1-Q)E\{x_2^2\}}{-\sum_{i=1}^M (d_i^2 - b_i^2)P_i} \quad (30)$$

4. Значение  $\hat{E}\{s_i^2(k)\}$  устанавливается в настроенную  $P_{Si}$ , и вычисляются веса  $w_{11}$ ,  $w_{12}$ ,  $w_{21}$  и  $w_{22}$

#### *В. Выбор между использованием четырех или двух весов*

Для многих случаев, два веса [18] достаточны для вычисления левого и правого поддиапазонов повторно микшированного сигнала [9]. В некоторых случаях лучшие результаты могут достигаться использованием четырех весов [13] и [15].

Использование двух весов означает, что, для формирования левого выходного сигнала используется только левый исходный сигнал и то же самое для правого выходного сигнала. Таким образом, сценарий, где желательны четыре веса, имеет место тогда, когда объект по одну сторону повторно микшируется, чтобы быть по другую сторону. В этом случае ожидалось бы, что использование четырех весов является благоприятным, так как сигнал, который изначально был только по одну сторону (например, в левом канале), по большей части будет по другую сторону (например, в правом канале) после повторного микширования. Таким образом, четыре веса могут использоваться для предоставления сигналу возможности перетекать из исходного левого канала в повторно микшированный правый канал, и наоборот.

Когда задача метода наименьших квадратов по вычислению четырех весов является плохо обусловленной, амплитуда весов может быть большой. Подобным образом, когда используется описанное выше повторное микширование с одной стороны на другую сторону, может быть большой амплитуда весов, когда используются только два веса. Оправданный этим наблюдением в некоторых реализациях может использоваться следующий критерий, чтобы решать, следует ли использовать четыре или два веса.

Если  $A < B$ , то используем четыре веса, иначе используем два веса.  $A$  и  $B$  являются показателями амплитуды весов для четырех и двух весов соответственно. В некоторых реализациях  $A$  и  $B$  вычисляются, как изложено ниже. Для вычисления  $A$  сначала

вычисляем четыре веса согласно [13] и [15], а затем, устанавливаем

$A = w_{11}^2 + w_{12}^2 + w_{21}^2 + w_{22}^2$ . Для вычисления  $B$ , веса могут вычисляться согласно [18], а затем вычисляется  $B = w_{11}^2 + w_{22}^2$ .

### 5 С. Улучшение степени затухания, когда требуется

Когда источник должен быть полностью удален, например с удалением дорожки ведущей вокальной партии для применения Караоке, его коэффициентами усиления микширования являются  $c_i=0$  и  $d_i=0$ . Однако когда пользователь выбирает нулевые коэффициенты усиления микширования, степень достижимого затухания может быть ограниченной. Таким образом, для улучшенного затухания значения мощностей поддиапазонов источника соответствующих сигналов источника, полученные из дополнительной информации,  $\hat{E} \{s_i^2(k)\}$ , могут масштабироваться значением большим, чем единица (например, 2), перед использованием для вычисления весов  $w_{11}$ ,  $w_{12}$ ,  $w_{21}$  и  $w_{22}$ .

### 15 D. Улучшение качества звукозаписи сглаживанием весов

Было обнаружено, что раскрытая схема повторного микширования может приносить артефакты в требуемый сигнал, особенно когда звуковой сигнал является тональным или стационарным. Чтобы улучшить качество звукозаписи, в каждом поддиапазоне может вычисляться показатель стационарности/тональности. Если показатель стационарности/тональности превышает определенное пороговое значение,  $TON_0$ , то веса оценки сглаживаются по времени. Операция сглаживания описана, как изложено ниже: для каждого поддиапазона, на каждом индексе  $k$  времени, получают веса, которые применяются для вычисления выходных поддиапазонов, как изложено ниже:

Если  $TON(k) > TON_0$ , то

$$\tilde{w}_{11}(k) = \alpha w_{11}(k) + (1 - \alpha) \tilde{w}_{11}(k-1),$$

$$30 \quad \tilde{w}_{12}(k) = \alpha w_{12}(k) + (1 - \alpha) \tilde{w}_{12}(k-1),$$

$$\tilde{w}_{21}(k) = \alpha w_{21}(k) + (1 - \alpha) \tilde{w}_{21}(k-1),$$

$$\tilde{w}_{22}(k) = \alpha w_{22}(k) + (1 - \alpha) \tilde{w}_{22}(k-1), \quad (31)$$

где  $\tilde{w}_{11}(k)$ ,  $\tilde{w}_{12}(k)$ ,  $\tilde{w}_{21}(k)$  и  $\tilde{w}_{22}(k)$  - сглаженные веса, а  $w_{11}(k)$ ,  $w_{12}(k)$ ,  $w_{21}(k)$  и  $w_{22}(k)$  - несглаженные веса, вычисляемые, как описано раньше.

Иначе

$$\tilde{w}_{11}(k) = w_{11}(k),$$

$$\tilde{w}_{12}(k) = w_{12}(k), \quad (32)$$

$$40 \quad \tilde{w}_{21}(k) = w_{21}(k),$$

$$\tilde{w}_{22}(k) = w_{22}(k).$$

### E. Регулирование окружения/реверберации

Технология повторного микширования, описанная в материалах настоящей заявки, предусматривает пользовательское управление в показателях коэффициентов  $c_i$  и  $d_i$  усиления микширования. Это соответствует определению для каждого объекта, коэффициента усиления,  $G_i$ , и амплитудного панорамирования,  $L_i$  (направления), где усиление и панорамирование полностью определяются посредством  $c_i$  и  $d_i$ :

$$50 \quad G_i = 10 \log_{10}(c_i^2 + d_i^2),$$

$$L_i = 20 \log_{10} \frac{c_i}{d_i}. \quad (33)$$

В некоторых реализациях может быть желательным регулировать другие признаки стереофонического микшированного сигнала, иные, чем усиление и амплитудное

панорамирование сигналов источников. В последующем описании описана технология для модификации степени окружения стереофонического звукового сигнала. Никакая дополнительная информация не используется для этой задачи декодера.

В некоторых реализациях модель прохождения сигналов, заданная в [44], может использоваться для модифицирования степени окружения стереофонического сигнала, где мощности поддиапазона у  $n_1$  и  $n_2$  предполагаются равными, то есть

$$E\{n_1^2(k)\} = E\{n_2^2(k)\} = P_N(k). \quad (34)$$

Вновь может быть допущено, что  $s$ ,  $n_1$  и  $n_2$  являются взаимно независимыми. При условии этих допущений, когерентность [17] может быть записана в виде:

$$\phi(k) = \frac{\sqrt{(E\{x_1^2(k)\} - P_N(k))(E\{x_2^2(k)\} - P_N(k))}}{\sqrt{E\{x_1^2(k)\}E\{x_2^2(k)\}}}. \quad (35)$$

Это соответствует квадратному уравнению с переменной  $P_N(k)$ :

$$P_N^2(k) - (E\{x_1^2(k)\} + E\{x_2^2(k)\})P_N(k) + E\{x_1^2(k)\}E\{x_2^2(k)\}(1 - \phi(k)^2) = 0. \quad (36)$$

Решениями этого квадратного уравнения являются:

$$P_N(k) = \frac{(E\{x_1^2(k)\} + E\{x_2^2(k)\}) \pm \sqrt{(E\{x_1^2(k)\} + E\{x_2^2(k)\})^2 - 4E\{x_1^2(k)\}E\{x_2^2(k)\}(1 - \phi(k)^2)}}{2}. \quad (37)$$

Физически возможным решением является такое с отрицательным знаком перед квадратным корнем:

$$P_N(k) = \frac{(E\{x_1^2(k)\} + E\{x_2^2(k)\}) - \sqrt{(E\{x_1^2(k)\} + E\{x_2^2(k)\})^2 - 4E\{x_1^2(k)\}E\{x_2^2(k)\}(1 - \phi(k)^2)}}{2}, \quad (38)$$

так как  $P_N(k)$  должна быть меньше, чем или равной  $E\{x_1^2(k)\} + E\{x_2^2(k)\}$ .

В некоторых реализациях для регулирования левого и правого окружения технология повторного микширования может применяться касательно двух объектов:

Одним объектом является источник с индексом  $i_1$ , с мощностью  $E\{s_{i_1}^2(k)\} = P_N(k)$  поддиапазона по левую сторону, то есть,  $a_{i_1} = 1$  и  $b_{i_1} = 0$ ; другим объектом является источник с индексом  $i_2$ , с мощностью  $E\{s_{i_2}^2(k)\} = P_N(k)$  поддиапазона по правую сторону, то есть  $a_{i_2} = 0$  и  $b_{i_2} = 1$ . Для изменения величины окружения пользователь может выбирать  $c_{i_1} = d_{i_1} = 10^{g_a/20}$  и  $c_{i_2} = d_{i_2} = 0$ , где  $g_a$  - прирост окружения в дБ.

#### *F. Иная дополнительная информация*

В некоторых реализациях модифицированная или иная дополнительная информация может использоваться в раскрытой схеме повторного микширования, которая более эффективна в показателях скорости передачи битов. Например, в [24],  $A_i(k)$  может иметь произвольные значения. Также есть зависимость от уровня исходного сигнала  $s_i(n)$  источника. Таким образом, для получения дополнительной информации в требуемом диапазоне уровню входного сигнала источника понадобилось бы настраиваться. Чтобы избежать этой настройки и чтобы устранить зависимость дополнительной информации от уровня исходного сигнала источника, в некоторых реализациях мощность поддиапазона источника может не только нормализоваться относительно мощности поддиапазона стереофонического сигнала, как в [24], но также могут учитываться коэффициенты усиления микширования:

$$A_i(k) = 10 \log_{10} \frac{(a_i^2 + b_i^2)E\{s_i^2\}}{E\{x_1^2(k)\} + E\{x_2^2(k)\}}. \quad (39)$$

Это соответствует использованию, в качестве дополнительной информации мощности источника, содержащейся в стереофоническом сигнале (а не непосредственно мощности источника), нормализованной стереофоническим сигналом. В качестве альтернативы можно использовать нормализацию, подобную

этой:

$$A_i(k) = 10 \log_{10} \frac{E\{s_i^2(k)\}}{\frac{1}{a_i^2} E\{x_1^2(k)\} + \frac{1}{b_i^2} E\{x_2^2(k)\}}. \quad (40)$$

5 Эта дополнительная информация также более эффективна, поскольку  $A_i(k)$  может принимать только значения, меньшие или равные 0 дБ. Отметим, что [39] и [40] могут быть решены для мощности  $E\{s_i^2(k)\}$  поддиапазона.

### *Г. Стереофонические сигналы/объекты источника*

10 Схема повторного микширования, описанная в материалах настоящей заявки, может быть легко расширена для манипулирования стереофоническими сигналами источников. С ракурса дополнительной информации стереофонические сигналы источника обрабатываются подобно двум монофоническим сигналам источника: один микшируется в левый, а другой микшируется в правый. То есть левый канал  $i$   
15 источника имеет ненулевой левый коэффициент  $a_i$  усиления и нулевой правый коэффициент  $b_{i+1}$  усиления. Коэффициенты усиления,  $a_i$  и  $b_{i+1}$ , могут оцениваться с помощью [6]. Дополнительная информация может передаваться, как если бы стереофонический источник был двумя монофоническими источниками. Некоторой информации необходимо передаваться в декодер, чтобы указывать декодеру, какие источники являются монофоническими источниками, а какие являются стереофоническими источниками.

Касательно обработки декодера и графического интерфейса пользователя (GUI) можно представлять в декодере стереофонический сигнал источника так же, как  
25 монофонический сигнал источника. То есть стереофонический сигнал источника имеет регулировку усиления и панорамирования, подобную монофоническому сигналу источника. В некоторых реализациях зависимость между регулировкой усиления и панорамирования GUI немикшированного повторно стереофонического сигнала и коэффициентами усиления может выбираться, чтобы быть:

$$30 \text{ GAIN}_0 = 0 \text{ dB},$$

$$\text{PAN}_0 = 20 \log_{10} \frac{b_{i+1}}{a_i}.$$

35 То есть GUI может устанавливаться в эти значения изначально. Зависимость между GAIN и PAN, выбранная пользователем, и новые коэффициенты усиления могут быть выбраны, чтобы быть:

$$\text{GAIN} = 10 \log_{10} \frac{(c_i^2 + d_{i+1}^2)}{(a_i^2 + b_{i+1}^2)},$$

$$40 \text{ PAN} = 20 \log_{10} \frac{d_{i+1}}{c_i}.$$

Уравнения [42] могут быть решены для  $c_i$  и  $d_{i+1}$ , которые могут использоваться в качестве коэффициентов усиления микширования (причем  $c_{i+1}=0$  и  $d_i=0$ ). Описанные функциональные возможности подобны регулировке «баланса» на стереофоническом усилителе. Усиления левого и правого каналов сигнала источника модифицируются без привнесения взаимного влияния.

## VI. Слепое формирование дополнительной информации

### *А. Полностью слепое формирование дополнительной информации*

50 В раскрытой схеме повторного микширования кодировщик принимает стереофонический сигнал и некоторое количество сигналов источников, представляющих объекты, которые должны повторно микшироваться в декодере. Дополнительная информация, необходимая для повторного микширования сигнала

источника с индексом  $i$  в декодере, определяется по коэффициентам  $a_i$  и  $b_i$  усиления и мощности  $E\{s_i^2(k)\}$  поддиапазона. Определение дополнительной информации описывалось в предыдущих разделах в случае, когда заданы сигналы источников.

5 Несмотря на то, что стереофонический сигнал легко получается (поскольку это соответствует изделию, существующему сегодня), может быть трудным получать сигналы источника, соответствующие объектам, которые должны повторно микшироваться в декодере. Таким образом, желательно формировать дополнительную информацию для повторного микширования, даже если сигналы источников объектов не доступны. В последующем описании описана технология полностью слепого формирования для формирования дополнительной информации только из стереофонического сигнала.

15 Фиг.8А - структурная схема реализации системы 800 кодирования, реализующей полностью слепое формирование дополнительной информации. Система 800 кодирования обычно включает в себя матрицу 802 гребенки фильтров, формирователь 804 дополнительной информации и кодировщик 806. Стереофонический сигнал принимается матрицей 802 гребенки фильтров, которая раскладывает стереофонический сигнал (например, левый и правый каналы) на пары поддиапазонов. Пары поддиапазонов принимаются процессором 804 дополнительной информации, который формирует дополнительную информацию из пар поддиапазонов с использованием требуемых перепада  $L_i$  уровней источника и функции  $f(M)$  усиления. Отметим, что ни матрица 802 гребенки фильтров, ни процессор 804 дополнительной информации не оперируют сигналами источников. 25 Дополнительная информация выводится полностью из входного стереофонического сигнала, требуемого перепада уровней источника,  $L_i$ , и функции усиления,  $f(M)$ .

Фиг.8В - блок-схема последовательности операций способа реализации последовательности 808 операций кодирования с использованием системы 800 кодирования по фиг.8А. Входной стереофонический сигнал раскладывается на пары поддиапазонов (810). Для каждого поддиапазона коэффициенты усиления,  $a_i$  и  $b_i$ , определяются для каждого требуемого сигнала источника с использованием требуемого значения перепада уровней источника,  $L_i$  (812). Для сигнала источника прямого звучания (например, сигнала источника, панорамированного по центру в павильоне звукозаписи), требуемым перепадом уровней источника является  $L_i=0$  дБ. При заданном  $L_i$  вычисляются коэффициенты усиления:

$$a_i = \frac{1}{\sqrt{1+A}},$$

(43)

$$b_i = \frac{\sqrt{A}}{\sqrt{1+A}},$$

45 где  $A=10^{L_i/10}$ . Отметим, что  $a_i$  и  $b_i$  были рассчитаны из условия, чтобы  $a_i^2+b_i^2=1$ . Это условие не является необходимостью, скорее оно является произвольным выбором для предохранения  $a_i$  или  $b_i$  от становления большими, когда велика амплитуда  $L_i$ .

Затем оценивается мощность поддиапазона прямого звучания с использованием пары поддиапазонов и коэффициентов усиления микширования (814). Для вычисления мощности поддиапазона прямого звучания можно допустить, что левый и правый поддиапазон каждого входного сигнала в каждый момент времени может быть выражен:

$$\begin{aligned}x_1 &= as + n_1, \\x_2 &= bs + n_2,\end{aligned}\tag{44}$$

где  $a$  и  $b$  - коэффициенты усиления микширования,  $s$  представляет прямое звучание всех сигналов источников, а  $n_1$  и  $n_2$  представляют независимое звуковое сопровождение. Может быть допущено, что  $a$  и  $b$  являются:

$$a = \frac{1}{\sqrt{1+B}},\tag{45}$$

$$b = \frac{\sqrt{B}}{\sqrt{1+B}},$$

где  $B = E\{x_2^2(k)\}/E\{x_1^2(k)\}$ . Отметим, что  $a$  и  $b$  могут вычисляться из условия, чтобы перепад уровней, с которым  $s$  содержится в  $x_2$  и  $x_1$ , являлся таким же, как перепад уровней между  $x_2$  и  $x_1$ . Перепадом уровней в дБ у прямого звучания является  $M = \log_{10} B$ .

Можем вычислять мощность поддиапазона прямого звучания,  $E\{s^2(k)\}$ , согласно модели прохождения сигналов, приведенной в [44]. В некоторых реализациях используется следующая система Уравнений:

$$E\{x_1^2(k)\} = a^2 E\{s^2(k)\} + E\{n_1^2(k)\},\tag{46}$$

$$E\{x_2^2(k)\} = b^2 E\{s^2(k)\} + E\{n_2^2(k)\},$$

$$E\{x_1(k)x_2(k)\} = abE\{s^2(k)\}.$$

Было допущено в [46], что  $s$ ,  $n_1$  и  $n_2$  в [34] взаимно независимы, левосторонние параметры в [46] могут измеряться, и доступны  $a$  и  $b$ . Таким образом, тремя неизвестными в [46] являются  $E\{s^2(k)\}$ ,  $E\{n_1^2(k)\}$  и  $E\{n_2^2(k)\}$ . Мощность поддиапазона прямого звучания,  $E\{s^2(k)\}$ , может быть задана посредством

$$E\{s^2(k)\} = \frac{E\{x_1(k)x_2(k)\}}{ab}.\tag{47}$$

Мощность поддиапазона прямого звучания также может быть записана в качестве функции когерентности [17]

$$E\{s^2(k)\} = \frac{\phi \sqrt{E\{x_1^2(k)\}E\{x_2^2(k)\}}}{ab}.\tag{48}$$

В некоторых реализациях вычисление требуемой мощности поддиапазона источника,  $E\{s_i^2(k)\}$ , может выполняться в два этапа: сначала вычисляется мощность поддиапазона прямого звучания,  $E\{s^2(k)\}$ , где  $s$  представляет прямое звучание всех источников (например, панорамированное по центру) в [44]; затем, вычисляются (816) требуемые мощности поддиапазонов источников,  $E\{s_i^2(k)\}$ , посредством модифицирования мощности поддиапазона прямого звучания,  $E\{s^2(k)\}$ , в качестве функции направления прямого звучания (представленного  $M$ ) и требуемого направления звучания (представленного требуемым перепадом  $L$  уровней источников):

$$E\{s_i^2(k)\} = f(M(k))E\{s^2(k)\},\tag{49}$$

где  $f(\cdot)$  - функция усиления, которая является функцией направления, возвращает коэффициент усиления, который близок к единице только для направления требуемого источника. В качестве заключительного этапа коэффициенты усиления и мощности  $E\{s_i^2(k)\}$  поддиапазонов могут квантоваться и кодироваться для формирования дополнительной информации (818).

Фиг.9 иллюстрирует примерную функцию  $f(M)$  усиления для требуемого перепада уровней источника,  $L_i = L$  dB. Отметим, что степень направленности может

регулироваться в показателях выбора  $f(M)$ , чтобы имела более или менее узкий пик вокруг требуемого направления  $L_0$ . Для требуемого источника в центре может использоваться ширина пика  $L_0=6$  дБ.

Отметим, что с полностью слепой технологией, описанной выше, может быть определена дополнительная информация ( $a_i, b_i, E\{s_i^2(k)\}$ ) для заданного сигнала  $s_i$  источника.

### *В. Сочетание между слепым и неслепым формированием дополнительной информации*

Технология полностью слепого формирования, описанная выше, может быть ограничена при определенных условиях. Например, если два объекта имеют одинаковое местоположение (направление) в павильоне стереофонической звукозаписи, то может не быть возможным вслепую формировать дополнительную информацию, относящуюся к одному или обоим объектам.

Альтернативой полностью слепому формированию дополнительной информации является частично слепое формирование дополнительной информации. Частично слепая технология формирует колебательный сигнал объекта, который грубо соответствует исходному колебательному сигналу объекта. Это, например, может делаться вынуждением певцов или музыкантов проигрывать/воспроизводить отдельный объектный сигнал. Либо можно применять данные MIDI (цифрового интерфейса музыкальных инструментов) для этой цели, или позволить синтезатору формировать объектный сигнал. В некоторых реализациях «необработанный» колебательный сигнал объекта является выравненным по времени со стереофоническим сигналом, относительно которого должна формироваться дополнительная информация. В таком случае дополнительная информация может формироваться с использованием последовательности операций, которая является комбинацией слепого и неслепого формирования дополнительной информации.

Фиг.10 - схема реализации последовательности 1000 операций формирования дополнительной информации с использованием технологии частично слепого формирования. Последовательность 1000 операций начинается получением входного стереофонического сигнала и  $M$  «необработанных» сигналов источников (1002). Затем коэффициенты  $a_i$  и  $b_i$  усиления определяются для  $M$  «необработанных» сигналов источников (1004). В каждом временном интервале, в каждом поддиапазоне, определяется первая краткосрочная оценка мощности поддиапазона,  $E\{s_i^2(k)\}$ , для каждого «необработанного» сигнала источника (1006). Вторая краткосрочная оценка мощности поддиапазона,  $\hat{E}\{s_i^2(k)\}$ , определяется для каждого «необработанного» сигнала источника с использованием технологии полностью слепого формирования, примененной к входному стереофоническому сигналу (1008).

В заключение к оцененным мощностям поддиапазонов применяется функция, которая объединяет первую и вторую оценки мощности поддиапазона и возвращает заключительную оценку, которая может эффективно использоваться для вычисления дополнительной информации (1010). В некоторых реализациях функция  $F()$  задана согласно

$$F(E\{s_i^2(k)\}, \hat{E}\{s_i^2(k)\}) \quad (50)$$

$$F(E\{s_i^2(k)\}, \hat{E}\{s_i^2(k)\}) = \min(E\{s_i^2(k)\}, \hat{E}\{s_i^2(k)\}).$$

VI. Архитектуры, пользовательские интерфейсы, синтаксис битового потока  
 А. Клиент/серверная архитектура

Фиг.11 - структурная схема реализации клиент-серверной архитектуры 1100 для поставки стереофонических сигналов и *M* сигналов источников и/или дополнительной информации в звуковые устройства 1110 с возможностью повторного микширования. Архитектура 1100 является только примером. Возможны другие архитектуры, в том числе архитектуры с большим или меньшим количеством компонентов.

Архитектура 1100, как правило, включает в себя службу 1102 загрузки, имеющую хранилище 1104 (например, MySQL™) и сервер 1106 (например, сервер Windows™ NT, Linux). Хранилище 1104 может хранить различные типы контента, включая профессионально микшированные стереофонические сигналы и ассоциативно связанные сигналы источников, соответствующие объектам в стереофонических сигналах и различным эффектам (например, реверберации). Стереофонические сигналы могут храниться в многообразии стандартизованных форматов, в том числе MP3, PCM, AAC и т.п.

В некоторых реализациях сигналы источников хранятся в хранилище 1104 и сделаны доступными для загрузки в звуковые устройства 1110. В некоторых реализациях предварительно обработанная дополнительная информация хранится в хранилище 1104 и сделана доступной для загрузки в звуковые устройства 1110. Предварительно обработанная дополнительная информация может формироваться сервером 1106 с использованием одной или более схем кодирования, описанных со ссылкой на фиг.1А, 6А и 8А.

В некоторых реализациях служба 1102 загрузки (например, веб-сайт, музыкальный магазин) поддерживает связь с звуковыми устройствами 1110 через сеть 1108 (например, сеть Интернет, сеть интранет, сеть Ethernet, беспроводную сеть, одноранговую сеть). Звуковые устройства 1110 могут быть любым устройством, допускающим реализацию раскрытых схем повторного микширования (например, медиаплеерами/устройствами записи, мобильными телефонами, персональными цифровыми секретарями (PDA), игровыми консолями, телевизионными абонентскими приставками, телевизионными приемниками, медиацентрами и т.п.).

#### *В. Архитектура звукового устройства*

В некоторых реализациях звуковое устройство 1110 включает в себя один или более процессоров или процессорных ядер 1112, устройства 1114 ввода (например, координатное колесо, мышь, джойстик, сенсорный экран), устройства 1120 вывода (например, ЖКД (жидкокристаллический дисплей LCD)), сетевые интерфейсы 1118 (например, USB, FireWire, Ethernet, плату сетевого интерфейса, беспроводный приемопередатчик) и машиночитаемый носитель 1116 (например, память, жесткий диск, флэш-память). Некоторые или все из этих компонентов могут отправлять и/или принимать информацию через каналы 1122 связи (например, шину, мост).

В некоторых реализациях машиночитаемый носитель 1116 включает в себя операционную систему, музыкальный диспетчер, звуковой процессор, модуль повторного микширования и музыкальную библиотеку. Операционная система ответственна за управление основными административными и коммуникационными задачами звукового устройства 1110, в том числе управление файлами, доступ к памяти, конфликты на шине, управление периферией, управление пользовательским интерфейсом, управление питанием и т.п. Музыкальный диспетчер может быть приложением, которое управляет музыкальной библиотекой. Звуковой процессор может быть традиционным звуковым процессором для воспроизведения музыкальных файлов (например, MP3, звукозаписи CD, и т.п.). Модуль повторного микширования может быть одним или более компонентов программного обеспечения, которые

реализуют функциональные возможности схем повторного микширования, описанных со ссылкой на фиг.1-10.

В некоторых реализациях сервер 1106 кодирует стереофонический сигнал и формирует дополнительную информацию, как описано в ссылках на фиг.1А, 6А и 8А.  
5 Стереофонический сигнал и дополнительная информация загружаются в звуковое устройство 1110 через сеть 1108. Модуль повторного микширования декодирует сигналы и дополнительную информацию и предоставляет возможность повторного микширования на основании пользовательского ввода, принятого через  
10 устройство 1114 ввода (например, клавиатуру, координатное колесо, сенсорный экран).

#### *С. Пользовательский интерфейс для приема пользовательского ввода*

Фиг.12 иллюстрирует реализацию пользовательского интерфейса 1202 для медиаплеера 1200 с возможностью повторного микширования. Пользовательский  
15 интерфейс 1202 также может быть адаптирован к другим устройствам (например, мобильным телефонам, компьютерам и т. п.). Пользовательский интерфейс не ограничен показанными конфигурацией и форматом и может включать в себя разные типы элементов пользовательского интерфейса (например, элементы управления навигацией, сенсорные интерфейсы).  
20

Пользователь может вводить режим «повторного микширования» для устройства 1200, выделяя надлежащий элемент на пользовательском интерфейсе 1202. В этом примере предполагается, что пользователь выбрал песню из музыкальной библиотеки, и пожелал бы изменить настройку панорамирования дорожки ведущей  
25 вокальной партии. Например, пользователь может пожелать слышать больше ведущей вокальной партии в левом звуковом канале.

Чтобы получить доступ к требуемому элементу управления панорамированием, пользователь может осуществлять навигацию по последовательности подменю 1204,  
30 1206 и 1208. Например, пользователь может осуществлять прокрутку для просмотра по пунктам в подменю 1204, 1206 и 1208 с использованием колеса 1210. Пользователь может выбирать выделенный пункт меню щелчком по кнопке 1212. Подменю 1208 предоставляет доступ к требуемому элементу управления панорамированием для дорожки ведущей вокальной партии. Затем пользователь может манипулировать  
35 ползунком (например, с использованием колеса 1210) для настройки панорамирования ведущей вокальной партии, как требуется, в то время как проигрывается песня.

#### *Д. Синтаксис битового потока*

В некоторых реализациях схемы повторного микширования, описанные со ссылкой  
40 на фиг.1-10, могут быть включены в существующие или будущие стандарты звукового кодирования (например, MPEG-4). Синтаксис битового потока для существующих и будущих стандартов кодирования может включать в себя информацию, которая может использоваться декодером с возможностью повторного микширования для  
45 определения, каким образом обрабатывать битовый поток, чтобы предусмотреть повторное микширование пользователем. Такой синтаксис может быть разработан для обеспечения обратной совместимости с традиционными схемами кодирования. Например, структура данных (например, заголовок пакета), включенный в битовый  
50 поток, может включать в себя информацию (например, один или более битов или флажковых признаков), указывающую наличие дополнительной информации (например, коэффициентов усиления, мощности поддиапазонов) для повторного микширования.

Раскрытые и другие варианты осуществления и функциональные операции, описанные в этом описании изобретения, могут быть реализованы в цифровой электронной схеме, или в компьютерном программном обеспечении, аппаратно реализованном программном обеспечении, включающих в себя структуры, раскрытие  
5 в этом описании изобретения и их структурные эквиваленты, или в комбинациях одного или более из них. Раскрытые и другие варианты осуществления могут быть реализованы в качестве одного или более компьютерных программных изделий, то есть одного или более модулей команд компьютерной программы, закодированных  
10 на машиночитаемом носителе для выполнения посредством или для управления работой устройства обработки данных. Машиночитаемый носитель может быть машиночитаемым устройством хранения данных, машиночитаемой запоминающей подложкой, устройством памяти, композицией, воздействующей на машиночитаемый распространяемый сигнал, или комбинацией одного или более из них. Термин  
15 «устройство обработки данных» охватывает все устройства, приспособления и машины для обработки данных, в том числе в качестве примера программируемый процессор, компьютер, либо многочисленные процессоры или компьютеры. Устройство может включать в себя в дополнение к аппаратным средствам машинную  
20 программу, которая создает среду выполнения для данной компьютерной программы, например машинную программу, которая составляет аппаратно реализованное программное обеспечение процессора, стек протоколов, систему управления базой данных, операционную систему или комбинацию одного или более из них. Распространяемый сигнал является искусственно сформированным сигналом,  
25 например сформированным машиной электрическим, оптическим или электромагнитным сигналом, который формируется, чтобы кодировать информацию для передачи на пригодное устройство приемника.

Компьютерная программа (также известная как программа, программное  
30 обеспечение, программно реализованное приложение, сценарий или код) может быть написана на любой разновидности языка программирования, в том числе компилируемых или интерпретируемых языках, и она может применяться в любой форме, в том числе в качестве автономной программы или в качестве модуля,  
35 компонента, процедуры или другого блока, пригодного для использования в вычислительной среде. Компьютерная программа не обязательно соответствует файлу или файловой системе. Программа может храниться в части файла, который удерживает другие программы или данные (например, один или более сценариев,  
40 хранимых в документе на языке разметки), в одиночном файле, выделенном для данной программы, или многочисленных скоординированных файлах (например, файлах, которые хранят один или более модулей, подпрограмм, или порций кода). Компьютерная программа может быть развернута, чтобы выполняться на одном компьютере или на многочисленных компьютерах, которые расположены на одной площадке или распределены по многочисленным площадкам и взаимосвязаны сетью  
45 связи.

Последовательности операций и логические потоки, описанные в этом описании изобретения, могут выполняться одним или более программируемых процессоров, приводящих в исполнение одну или более компьютерных программ для выполнения  
50 функций посредством оперирования с входными данными и формирования выходных сигналов. Последовательности операций и логические потоки также могут выполняться посредством, и устройство также может быть реализовано в качестве логической схемы специального назначения, например FPGA (программируемой

пользователем вентиляционной матрицы) или ASIC (специализированной интегральной схемы).

Процессоры, пригодные для выполнения компьютерной программы, включают в себя в качестве примера микропроцессоры как общего применения, так и  
5 специального назначения, и любые один или более процессоров любой разновидности цифрового компьютера. Обычно процессор будет принимать команды и данные из постоянного запоминающего устройства или оперативного запоминающего устройств, либо обоих. Существенными элементами компьютера являются процессор  
10 для выполнения команд и одно или более устройств памяти для хранения команд и данных. Как правило, компьютер также будет включать в себя или будет оперативно присоединен для приема данных из или передачи данных в, либо того и другого к одному или более устройствам хранения большой емкости для хранения данных, например магнитным, магнитооптическим дискам или оптическим дискам. Однако  
15 компьютеру не обязательно содержать такие устройства. Машиночитаемые носители для хранения команд и данных компьютерной программы включают в себя все разновидности энергозависимой памяти, носителей и устройств памяти, в том числе в качестве примера полупроводниковые устройства памяти, например СППЗУ (стираемое программируемое постоянное запоминающее устройство, EPROM), ЭСППЗУ (электрически стираемое программируемое постоянное запоминающее устройство, EEPROM), и устройства флэш-памяти; магнитные диски, например  
20 внутренние жесткие диски или съемные диски; магнитооптические диски; и диски CD-ROM (ПЗУ на компакт диске) и DVD-ROM (ПЗУ на многофункциональном цифровом диске). Процессор и память могут быть дополнены или включены в состав логической схемой специального назначения.

Чтобы предусмотреть взаимодействие с пользователем, раскрытые варианты осуществления могут быть реализованы на компьютере, имеющем в распоряжении  
30 устройство отображения, например монитор с ЭЛТ (электронно-лучевой трубкой, CRT), либо ЖКД (жидкокристаллическим дисплеем), для отображения информации пользователю, а также клавиатуру и координатно-указательное устройство, например мышь или шаровой манипулятор, посредством которых пользователь может выдавать входные данные в компьютер. Другие виды устройств могут  
35 использоваться, чтобы также предусматривать взаимодействие с пользователем; например, обратная связь, предоставляемая пользователю, может быть любой формой сенсорной обратной связи, например визуальной обратной связью, слуховой обратной связью, тактильной обратной связью; и входные данные от пользователя  
40 могут приниматься в любой форме, включая акустический, речевой или тактильный ввод.

Раскрытые варианты осуществления могут быть реализованы в вычислительной системе, которая включает в себя компонент прикладного программного обеспечения, например как сервер данных, или которая включает в себя компонент  
45 межплатформенного программного обеспечения, например сервер приложений, или которая включает в себя компонент интерфейсного программного обеспечения, например клиентский компьютер, имеющий графический интерфейс пользователя или веб-браузер, с помощью которого пользователь может взаимодействовать с реализацией того, что здесь описано, или любое сочетание одного или более  
50 компонентов прикладного программного обеспечения, межплатформенного программного обеспечения или интерфейсного программного обеспечения. Компоненты системы могут быть взаимосвязаны любой формой или носителем

цифровой передачи данных, например сетью связи. Примеры сетей связи включают в себя локальную сеть («LAN») и глобальную сеть («WAN»), например сеть Интернет.

Вычислительная система может включать в себя клиентов и серверы. Клиент и сервер обычно являются удаленными друг от друга и типично взаимодействуют через сеть связи. Взаимосвязь клиента и сервера возникает в силу компьютерных программ, работающих на соответствующих компьютерах и имеющих клиент-серверное взаимоотношение друг с другом.

Vii. Примеры систем, использующих технологию повторного микширования

Фиг.13 иллюстрирует реализацию системы 1300 декодера, объединяющей декодирование пространственно кодированных звуковых объектов (SAOC) и декодирование повторного микширования. SAOC является технологией звукозаписи для обработки многоканального аудио, которая предоставляет возможность интерактивного управления кодированными объектами звучания.

В некоторых реализациях система 1300 включает в себя декодер 1301 сигнала микширования, формирователь 1302 параметров и рендерер 1304 повторного микширования. Формирователь 1302 параметров включает в себя блок 1308 слепой оценки, формирователь 1310 пользовательских параметров микширования и формирователь 1306 параметров повторного микширования. Формирователь 1306 параметров повторного микширования включает в себя формирователь 1312 параметров эквивалентного микширования и формирователь 1314 параметров повышающего микширования.

В некоторых реализациях система 1300 предусматривает две последовательности операций звуковой обработки. В первой последовательности операций дополнительная информация, поставляемая системой кодирования, используется формирователем 1306 параметров повторного микширования для формирования параметров повторного микширования. Во второй последовательности операций слепые параметры формируются блоком 1308 слепой оценки и используются формирователем 1306 параметров повторного микширования для формирования параметров повторного микширования. Слепые параметры и последовательности операций полностью или частично слепого формирования могут выполняться блоком 1308 слепой оценки, как описано со ссылкой на фиг.8А и 8В.

В некоторых реализациях формирователь 1306 параметров повторного микширования принимает дополнительную информацию или слепые параметры и набор пользовательских параметров микширования из формирователя 1310 пользовательских параметров микширования. Формирователь 1310 пользовательских параметров микширования принимает параметры микширования, заданные конечными пользователями (например, усиление, панорамирование), и преобразует параметры микширования в формат, пригодный для обработки повторного микширования формирователем 1306 параметров повторного микширования (например, преобразует в коэффициенты  $c_i$ ,  $d_{i+1}$  усиления). В некоторых реализациях формирователь 1310 пользовательских параметров микширования дает пользователю интерфейс для предоставления пользователю возможности задавать требуемые параметры микширования, например, такой как пользовательский интерфейс 1200 медиаплеера, как описано со ссылкой на фиг.12.

В некоторых реализациях формирователь 1306 параметров микширования может обрабатывать как стереофонические, так и многоканальные звуковые сигналы. Например, формирователь 1312 параметров эквивалентного микширования может формировать параметры повторного микширования для целевого сигнала

стереофонических каналов, а формирователь 1314 параметров повышающего микширования может формировать параметры повторного микширования для многоканального целевого сигнала. Формирователь параметров повторного микширования, основанный на многоканальных звуковых сигналах, описывался со ссылкой на раздел IV.

В некоторых реализациях рендерер 1304 повторного микширования принимает параметры повторного микширования для стереофонического целевого сигнала или многоканального целевого сигнала. Рендерер 1316 эквивалентного микширования применяет параметры стереофонического повторного микширования к исходному стереофоническому сигналу, принятому непосредственно из декодера 1301 сигнала микширования, чтобы выдавать требуемый повторно микшированный стереофонический сигнал на основании форматированных заданных пользователем параметров стереофонического микширования, поставляемых формирователем 1310 пользовательских параметров микширования. В некоторых реализациях параметры стереофонического микширования могут применяться к исходному стереофоническому сигналу с использованием матрицы  $n \times n$  (например, матрицы  $2 \times 2$ ) параметров стереофонического микширования. Рендерер 1318 повышающего микширования применяет параметры многоканального повторного микширования к исходному многоканальному сигналу, принятому непосредственно из декодера 1301 сигнала микширования, чтобы выдавать требуемый повторно микшированный многоканальный сигнал на основании форматированных заданных пользователем параметров многоканального микширования, поставляемых формирователем 1310 пользовательских параметров микширования. В некоторых реализациях формирователь 1320 эффектов формирует сигналы эффектов (например, реверберацию), которые должны применяться к исходным стереофоническим или многоканальным сигналам рендерером 1316 эквивалентного микширования или рендерером повышающего микширования соответственно. В некоторых реализациях рендерер 1318 повышающего микширования принимает исходный стереофонический сигнал и преобразует (или осуществляет повышающее микширование) стереофонический сигнал в многоканальный сигнал в дополнение к применению параметров повторного микширования, чтобы сформировать повторно микшированный многоканальный сигнал.

Система 1300 может обрабатывать звуковые сигналы, имеющие многообразие конфигураций каналов, предоставляющих системе 1300 возможность интегрироваться в существующие схемы звукового кодирования (например, SAOC, MPEG AAC, параметрический стереофонический сигнал), наряду с сохранением обратной совместимости с такими схемами звукового кодирования.

Фиг.14А иллюстрирует обычную модель микширования для отдельного уровня громкости диалога (SDV). SDV является улучшенной технологией расширения диалога, описанной в предварительной заявке на выдачу патента США, под № 60/884,594 на «Separate Dialogue Volume». В одной из реализаций SDV стереофонические сигналы записываются и микшируются из условия, чтобы для каждого источника сигнал когерентно входил в левый и правый каналы сигнала с отдельными контрольными сигналами направления (например, перепадом уровней, разновременностью), а отраженные/реверберированные независимые сигналы входили в каналы, определяющие ширину слухового явления и контрольные сигналы охвата слушателя. Со ссылкой на фиг.14А коэффициент  $a$  определяет направление, в котором имеет место слуховое явление, где  $s$  - прямой звук, а  $n_1$  и  $n_2$  - боковые

отражения. Сигнал  $s$  имитирует локализованное звучание с направления, определенного коэффициентом  $a$ . Независимые сигналы,  $n_1$  и  $n_2$ , соответствуют отраженному/реверберированному звуку, часто указываемому звуковым сопровождением или окружением. Описанный сценарий является мотивированным восприятием разложением для стереофонических сигналов с одним звуковым источником:

$$\begin{aligned} x_1(n) &= s(n) + n_1 \\ x_2(n) &= as(n) + n_2, \end{aligned} \quad (51)$$

фиксирующим местоположение звукового источника и окружения.

Фиг.14В иллюстрирует реализацию системы 1400, объединяющей SDV с технологией повторного микширования. В некоторых реализациях система 1400 включает в себя гребенку 1402 фильтров (например, STFT), блок 1404 слепой оценки, рендерер 1406 эквивалентного микширования, формирователь 1408 параметров и обратную гребенку 1410 фильтров (например, обратное STFT).

В некоторых реализациях сигнал понижающего микширования SDV принимается и раскладывается гребенкой 1402 фильтров на сигналы поддиапазонов. Сигнал понижающего микширования может быть стереофоническим сигналом,  $x_1$ ,  $x_2$ , заданным согласно [51]. Сигналы  $X_1(i, k)$ ,  $X_2(i, k)$  поддиапазонов вводятся либо непосредственно в рендерер 1406 эквивалентного микширования либо в блок 1404 слепой оценки, который выводит слепые параметры,  $A$ ,  $P_S$ ,  $P_N$ . Вычисление этих параметров описано в предварительной заявке на выдачу патента США под № 60/884,594 на «Separate Dialogue Volume». Слепые параметры вводятся в формирователь 1408 параметров, который формирует параметры эквивалентного микширования,  $w_{11}$ –  $w_{22}$ , из слепых параметров и заданных пользователем параметров  $g(i, k)$  микширования (например, центрального усиления, ширины центра, частоты среза, сухости). Вычисление параметров эквивалентного микширования описано в разделе I. Параметры эквивалентного микширования применяются к сигналам поддиапазонов рендерером 1406 эквивалентного микширования, чтобы выдавать подвергшиеся рендерингу выходные сигналы,  $y_1$ ,  $y_2$ . Подвергшиеся рендерингу выходные сигналы рендерера 1406 эквивалентного микширования вводятся в обратную гребенку 1410 фильтров, которая преобразует подвергшиеся рендерингу выходные сигналы в требуемый стереофонический сигнал SDV на основании заданных пользователем параметров микширования.

В некоторых реализациях система 1400 также может обрабатывать звуковые сигналы с использованием технологии повторного микширования, как описано со ссылкой на фиг.1-12. В режиме повторного микширования гребенка 1402 фильтров принимает стереофонические или многоканальные сигналы, такие как сигналы, описанные в [1] и [27]. Сигналы раскладываются на сигналы  $X_1(i, k)$ ,  $X_2(i, k)$  поддиапазонов гребенкой 1402 фильтров и вводятся непосредственно в эквивалентный рендерер 1406 и блок 1404 слепой оценки для оценки слепых параметров. Слепые параметры вводятся в формирователь 1408 вместе с дополнительной информацией  $a_i$ ,  $b_i$ ,  $P_{s_i}$ , принятой в битовом потоке. Формирователь 1408 параметров применяет слепые параметры и дополнительную информацию к сигналам поддиапазонов, чтобы сформировать подвергшиеся рендерингу выходные сигналы. Подвергшиеся рендерингу выходные сигналы вводятся в обратную гребенку 1410 фильтров, которая формирует требуемый сигнал повторного микширования.

Фиг.15 иллюстрирует реализацию рендерера 1406 эквивалентного микширования, показанного на фиг.14В. В некоторых реализациях сигнал X1 понижающего микширования масштабируется модулями 1502 и 1504 масштабирования, а сигнал X2 понижающего микширования масштабируется модулями 1506 и 1508 масштабирования. Модуль 1502 масштабирования масштабирует сигнал X1 понижающего микширования параметром  $w_{11}$  эквивалентного микширования, модуль 1504 масштабирования масштабирует сигнал X1 понижающего микширования параметром  $w_{21}$  эквивалентного микширования, модуль 1506 масштабирования масштабирует сигнал X2 понижающего микширования параметром  $w_{12}$  эквивалентного микширования, и модуль 1508 масштабирования масштабирует сигнал X2 понижающего микширования параметром  $w_{22}$  эквивалентного микширования. Выходные сигналы модулей 1502 и 1506 масштабирования суммируются, чтобы выдавать первый подвергшийся рендерингу выходной сигнал  $u_1$ , а модулей 1504 и 1508 масштабирования суммируются, чтобы выдавать второй подвергшийся рендерингу выходной сигнал  $u_2$ .

Фиг.16 иллюстрирует систему 1600 распространения для технологии повторного микширования, описанной относительно фиг.1-15. В некоторых реализациях поставщик 1602 контента использует авторское инструментальное средство 1606, которое включает в себя кодировщик 1606 повторного микширования, для формирования дополнительной информации, как описано ранее относительно фиг.1А. Дополнительная информация может быть частью одного или более файлов и/или включаться в битовый поток для услуги потоковой передачи битов. Файлы повторного микширования могут иметь уникальное расширение имени файла (например, filename.gmx). Одиночный файл может включать в себя исходный микшированный звуковой сигнал и дополнительную информацию. В качестве альтернативы исходный микшированный звуковой сигнал и дополнительная информация могут распространяться в качестве отдельных файлов в пакете, комплекте, упаковке или другом подходящем контейнере. В некоторых реализациях файлы повторного микширования могут распространяться с предустановленными параметрами микширования, чтобы помогать пользователю изучать технологию и/или для маркетинговых целей.

В некоторых реализациях исходный контент (например, исходный файл микшированного аудио), дополнительная информация и необязательные предустановленные параметры микширования («информация повторного микширования») могут поставляться поставщику 1608 услуг (например, в музыкальный портал) или размещаться на физическом носителе (например, CD-ROM, DVD, медиаплеере, флэш-памяти). Поставщик 1608 услуг может управлять одним или более серверами 1610 для обслуживания всей или части информации повторного микширования и/или битовых потоков, содержащих всю или часть информации повторного микширования. Информация повторного микширования может храниться в хранилище 1612. Поставщик 1608 услуг также может предоставлять виртуальную среду (например, социальную общность, портал, доску объявлений) для совместного использования сформированных пользователем параметров микширования. Например, параметры микширования, сформированные пользователем в подготовленном для повторного микширования устройстве 1616 (например, медиаплеере, мобильном телефоне), могут сохраняться в файле параметров микширования, который может выгружаться поставщику 1608 услуг для совместного использования с другими пользователями. Файлы параметров

микширования могут иметь уникальное расширение имени (например, filename.rms). В показанном примере пользователь сформировал файл параметров микширования с использованием устройства А воспроизведения повторного микширования и выгрузил файл параметров микширования к поставщику 1608 услуг, где файл  
5 впоследствии загружался пользователем, эксплуатирующим устройство В воспроизведения повторного микширования.

Система 1600 может быть реализована с использованием любой известной схемы управления цифровыми правами и/или другими известными способами защиты для  
10 защиты исходного контента и информации повторного микширования. Например, пользователю, эксплуатирующему устройство В воспроизведения повторного микширования, может понадобиться загрузить исходный контент отдельно и защитить право использования до того, как пользователь может осуществлять доступ  
15 или пользоваться признаками повторного микширования, предоставленными устройством В воспроизведения повторного микширования.

Фиг.17А иллюстрирует основные элементы битового потока для предоставления информации повторного микширования. В некоторых реализациях на  
20 задействованное повторным микшированием устройство может доставляться единый интегрированный битовый поток 1702, который включает в себя микшированный звуковой сигнал (Mixed\_Obj BS), коэффициенты усиления и мощности поддиапазонов (Ref\_Mix\_Para BS), а также заданные пользователем параметры микширования (User\_Mix\_Para BS). В некоторых реализациях многочисленные битовые потоки для информации повторного микширования могут независимо  
25 доставляться на задействованные повторным микшированием устройства. Например, микшированный звуковой сигнал может доставляться в первом битовом потоке 1704, а коэффициенты усиления, мощности поддиапазонов и заданные пользователем параметры микширования могут доставляться во втором битовом потоке 1706. В  
30 некоторых реализациях микшированный звуковой сигнал, коэффициенты усиления и мощности поддиапазонов, и заданные пользователем параметры микширования могут доставляться в трех отдельных битовых потоках, 1708, 1710 и 1712. Эти отдельные битовые потоки могут доставляться на одной и той же или разных скоростях передачи битов. Битовые потоки могут обрабатываться по необходимости с  
35 использованием многообразия известных технологий для сбережения полосы пропускания и обеспечения надежности, в том числе побитовое перемежение, энтропийное кодирование (например, кодирование кодом Хаффмана), исправление ошибок и т.п.

Фиг.17В иллюстрирует интерфейс битового потока для кодировщика 1714 повторного микширования. В некоторых реализациях входные сигналы в  
40 интерфейс 1714 кодировщика повторного кодирования могут включать в себя микшированный объектный сигнал, индивидуальные сигналы объекта или источника и варианты выбора кодировщика. Выходные сигналы интерфейса 1714 кодировщика  
45 могут включать в себя битовый поток микшированного звукового сигнала, битовый поток, включающий в себя коэффициенты усиления и мощности поддиапазонов, и битовый поток, включающий в себя предустановленные параметры микширования.

Фиг.17С иллюстрирует интерфейс битового потока для декодера 1716 повторного микширования. В некоторых реализациях входные сигналы в интерфейс 1716 декодера  
50 могут включать в себя битовый поток микшированного звукового сигнала, битовый поток, включающий в себя коэффициенты усиления и мощности поддиапазонов, и битовый поток, включающий в себя предустановленные параметры микширования.

Выходные сигналы интерфейса 1716 декодера могут включать в себя повторно микшированный звуковой сигнал, битовый поток рендерера повышающего микширования (например, многоканальный сигнал), слепые параметры повторного микширования и пользовательские параметры повторного микширования.

5       Возможны другие конфигурации для интерфейсов кодировщика и декодера. Конфигурации интерфейсов, проиллюстрированные на фиг.17В и 17С, могут использоваться, чтобы определять интерфейс прикладного программирования (API) для предоставления задействованным повторным микшированием устройствам  
10 возможности обрабатывать информацию повторного микширования. Показанные интерфейсы, проиллюстрированные на фиг.17В и 17С, являются примерами, и возможны другие конфигурации, в том числе конфигурации с разными количествами и типами входных и выходных сигналов, которые могут быть частично основаны на устройстве.

15       Фиг.18 - структурная схема, показывающая примерную систему 1800, включающую в себя расширения для формирования добавочной дополнительной информации для определенных объектных сигналов, чтобы обеспечивать улучшенное воспринимаемое качество повторно микшированного сигнала. В некоторых реализациях система 1800  
20 включает в себя (на стороне кодирования) кодировщик 1808 сигнала микширования и усовершенствованный кодировщик 1802 повторного микширования, который включает в себя кодировщик 1804 повторного микширования и кодировщик 1806 сигналов. В некоторых реализациях система 1800 включает в себя (на стороне декодирования) декодер 1810 сигнала микширования, рендерер 1814 повторного  
25 микширования и формирователь 1816 параметров.

      На стороне кодировщика микшированный звуковой сигнал кодируется кодировщиком 1808 сигнала микширования (например, кодировщиком mp3) и отправляется на сторону декодирования. Объектные сигналы (например, ведущая  
30 вокальная партия, гитара, барабаны или другие инструменты) вводятся в кодировщик 1804 повторного микширования, который формирует дополнительную информацию (например, коэффициенты усиления и мощности поддиапазонов), например, как описано ранее со ссылкой на фиг.1А и 3А. Дополнительно один или  
35 более интересующих объектных сигналов вводятся в кодировщик 1806 сигналов (например, кодировщик mp3), чтобы вырабатывать добавочную дополнительную информацию. В некоторых реализациях информация выравнивания вводится в кодировщик 1806 сигналов для выравнивания выходных сигналов кодировщика 1808  
40 сигнала микширования и кодировщика 1806 сигналов соответственно. Информация выравнивания может включать в себя информацию временного выравнивания, тип используемого кодека, целевую скорость передачи битов, информацию или стратегию распределения битов и т.п.

      На стороне декодера выходной сигнал кодировщика сигнала микширования вводится в декодер 1810 сигнала микширования (например, декодер mp3). Выходной  
45 сигнал декодера 1810 сигнала микширования и дополнительная информация кодировщика (например, сформированные кодировщиком коэффициенты усиления, мощности поддиапазонов, добавочная дополнительная информация) вводятся в формирователь 1816 параметров, который использует эти параметры, вместе с  
50 параметрами управления (например, заданными пользователем параметрами микширования), для формирования параметров повторного микширования и дополнительных данных повторного микширования. Параметры повторного микширования и дополнительные данные повторного микширования могут

использоваться формирователем (рендерером) 1814 повторно микшированного сигнала для формирования повторно микшированного звукового сигнала.

Дополнительные данные повторного микширования (например, объектный сигнал) используются формирователем 1814 повторно микшированного сигнала для  
 5 повторного микширования конкретного объекта в исходном звуковом сигнале микширования. Например, в приложении караоке объектный сигнал, представляющий ведущую вокальную партию, может использоваться усовершенствованным кодировщиком 1802 повторного микширования для формирования добавочной  
 10 дополнительной информации (например, кодированного объектного сигнала). Этот сигнал может использоваться формирователем 1816 параметров для формирования дополнительных данных повторного микширования, которые могут использоваться формирователем 1814 повторно микшированного сигнала для повторного микширования ведущей вокальной партии в исходном звуковом сигнале  
 15 микширования (например, с подавлением или ослаблением ведущей вокальной партии).

Фиг.19 - структурная схема, показывающая пример формирователя 1814 повторно микшированного сигнала, показанного на фиг.18. В некоторых реализациях  
 20 сигналы X1, X2 повторного микширования вводятся в объединители 1904, 1906 соответственно. Сигналы X1, X2 понижающего микширования, например, могут быть левым и правым каналами исходного звукового сигнала микширования. Объединители 1904, 1906 объединяют сигналы X1, X2 понижающего микширования с  
 25 дополнительными данными повторного микширования, поставляемыми формирователем 1816 параметров. В примере караоке объединение может включать в себя вычитание объектного сигнала ведущей вокальной партии из сигналов X1, X2 понижающего микширования перед повторным микшированием для ослабления или подавления ведущей вокальной партии в повторно микшированном звуковом сигнале.

В некоторых реализациях сигнал X1 понижающего микширования (например,  
 30 левый канал исходного звукового сигнала микширования) объединяется с дополнительными данными повторного микширования (например, левым каналом объектного сигнала ведущей вокальной партии) и масштабируется модулями 1906а и 1906b масштабирования, а сигнал X2 понижающего микширования (например,  
 35 правый канал исходного звукового сигнала микширования) объединяется с дополнительными данными повторного микширования (например, правым каналом объектного сигнала ведущей вокальной партии) и масштабируется модулями 1906с и 1906d масштабирования. Модуль 1906а масштабирования масштабирует сигнал X1  
 40 понижающего микширования параметром  $w_{11}$  эквивалентного микширования, модуль 1906b масштабирования масштабирует сигнал X1 понижающего микширования параметром  $w_{21}$  эквивалентного микширования, модуль 1906с масштабирования масштабирует сигнал X2 понижающего микширования параметром  $w_{12}$  эквивалентного микширования, и модуль 1906d масштабирования масштабирует  
 45 сигнал X2 понижающего микширования параметром  $w_{22}$  эквивалентного микширования. Масштабирование может быть реализовано с использованием линейной алгебры, к примеру с использованием матрицы  $n$  на  $n$  (например,  $2 \times 2$ ). Выходные сигналы модулей 1906а и 1906с масштабирования суммируются, чтобы  
 50 выдавать первый подвергшийся рендерингу выходной сигнал Y1, а модулей 1906b и 1906d масштабирования суммируются, чтобы выдавать второй подвергшийся рендерингу выходной сигнал Y2.

В некоторых реализациях можно реализовать элемент управления (например,

переключатель, ползунок, кнопку) в пользовательском интерфейсе, чтобы переходить между исходным стереофоническим микшированием, режимом «караоке» и/или режимом «без инструментального сопровождения». В качестве функции положения этого управляющего элемента объединитель 1902 управляет линейной комбинацией между исходным стереофоническим сигналом и сигналом(ами), полученным посредством добавочной дополнительной информации. Например, для режима караоке сигнал, полученный из добавочной дополнительной информации, может вычитаться из стереофонического сигнала. Обработка повторным микшированием может применяться впоследствии для удаления шума квантования (если стереофонический и/или другой сигнал кодировался с потерями). Для частичного удаления вокальных партий необходимо вычитаться только части сигнала, полученного посредством добавочной дополнительной информации. Для воспроизведения только вокальных партий, объединитель 1902 выбирает сигнал, полученный посредством добавочной дополнительной информации. Для воспроизведения вокальных партий с некоторой фоновой музыкой объединитель 1902 добавляет масштабированный вариант стереофонического сигнала к сигналу, полученному посредством добавочной дополнительной информации.

Несмотря на то, что это описание изобретения содержит много особенностей, таковые должны интерпретироваться не в качестве ограничений на объем того, что является формулой изобретения, или того, что может быть заявлено формулой изобретения, а скорее в качестве описаний признаков, специфичных конкретным вариантам осуществления. Определенные признаки, которые описаны в этом описании изобретения в контексте отдельных вариантов осуществления, также могут быть реализованы в комбинации в едином варианте осуществления. Наоборот, различные признаки, которые описаны в контексте одиночного варианта осуществления, также могут быть реализованы в многочисленных вариантах осуществления, отдельно или в любой подкомбинации. Более того, хотя признаки могут быть описаны выше как действующие в определенных комбинациях, и даже изначально заявлены как таковые, один или более признаков из заявленной комбинации, в некоторых случаях, могут исключаться из комбинации, и заявленная комбинация может быть направлена на подкомбинацию или вариант подкомбинации.

Подобным образом, несмотря на то, что операции описаны на чертежах в конкретной очередности, это не должно пониматься в качестве требования, чтобы такие операции выполнялись в конкретной показанной очередности или в последовательном порядке, или чтобы выполнялись все проиллюстрированные операции для достижения требуемых результатов. В определенных случаях могут быть полезны многозадачность и параллельная обработка. Более того, разделение различных компонентов системы в вариантах осуществления, описанных выше, не должно пониматься в качестве требования такого разделения во всех вариантах осуществления, и должно быть понятно, что описанные программные компоненты и системы, в общем смысле, могут быть объединены вместе в едином программном обеспечении или упакованы в многочисленные программные изделия.

В этом описании изобретения были описаны конкретные варианты осуществления объекта изобретения. Другие варианты осуществления находятся в объеме последующей формулы изобретения. Например, действия, перечисленные в формуле изобретения, могут выполняться в разной очередности и по-прежнему достигать требуемых результатов. В качестве одного из примеров последовательность операций, изображенная на прилагаемых фигурах, не обязательно требует конкретной

показанной очередности, или последовательного порядка, для достижения требуемых результатов.

В качестве еще одного примера предварительная обработка дополнительной информации, описанная в разделе 5А, предусматривает нижнюю границу в мощности поддиапазона повторно микшированного сигнала для предотвращения отрицательных значений, которые вступают в противоречие с моделью прохождения сигналов, приведенной в [2]. Однако эта модель прохождения сигналов подразумевает не только положительную мощность повторно микшированного сигнала, но также положительные векторные произведения между исходными стереофоническими сигналами и повторно микшированными стереофоническими сигналами, а именно  $E\{x_1y_1\}$ ,  $E\{x_1y_2\}$ ,  $E\{x_2y_1\}$  и  $E\{x_2y_2\}$ .

Начиная со случая с двумя весами, чтобы не допустить, чтобы векторные произведения  $E\{x_1y_1\}$  и  $E\{x_2y_2\}$  становились отрицательными, веса, определенные в [18], ограничиваются определенным пороговым значением, из условия, чтобы они никогда не были меньшими, чем А дБ.

В таком случае векторные произведения ограничиваются, учитывая следующие условия, где  $\sqrt{\phantom{x}}$  обозначает квадратный корень, а  $Q$  определено в качестве  $Q=10^{A/10}$ :

Если  $E\{x_1y_1\} < Q^2 E\{x_1^2\}$ , то векторное произведение ограничено до  $E\{x_1y_1\} = Q^2 E\{x_1^2\}$ .

Если  $E\{y_1, y_2\} < Q \sqrt{E\{x_1^2\} E\{x_2^2\}}$ , то векторное произведение ограничено до  $E\{x_1y_2\} = Q \sqrt{E\{x_1^2\} E\{x_2^2\}}$ .

Если  $E\{x_2, y_1\} < Q \sqrt{E\{x_1^2\} E\{x_2^2\}}$ , то векторное произведение ограничено до  $E\{x_2y_1\} = Q \sqrt{E\{x_1^2\} E\{x_2^2\}}$ .

Если  $E\{x_2y_2\} < Q^2 E\{x_2^2\}$ , то векторное произведение ограничено до  $E\{x_2y_2\} = Q^2 E\{x_2^2\}$ .

#### Формула изобретения

1. Способ декодирования многоканального звукового сигнала с повторным микшированием, содержащий этапы, на которых:

получают исходный многоканальный звуковой сигнал, содержащий набор сигналов источников, который включает в себя один или более сигналов источников, предназначенных для повторного микширования;

получают дополнительную информацию, представляющую, по меньшей мере, одно из зависимости между исходным многоканальным звуковым сигналом и каждым из набора сигналов источников и энергетического уровня каждого из набора сигналов источников;

получают набор параметров микширования для повторного микширования исходного многоканального звукового сигнала и

формируют повторно микшированный многоканальный звуковой сигнал с использованием исходного многоканального звукового сигнала, дополнительной информации и набора параметров микширования.

2. Способ по п.1, в котором формирование повторно микшированного многоканального звукового сигнала содержит этапы, на которых:

раскладывают исходный многоканальный звуковой сигнал на первый набор сигналов поддиапазонов;

оценивают второй набор сигналов поддиапазонов, соответствующий повторно микшированному многоканальному звуковому сигналу, с использованием дополнительной информации и набора параметров микширования и преобразуют второй набор сигналов поддиапазонов в повторно микшированный многоканальный звуковой сигнал.

3. Способ по п.2, в котором оценивание второго набора сигналов поддиапазонов дополнительно содержит этапы, на которых:

декодируют дополнительную информацию, чтобы предоставить коэффициенты усиления и оценки мощностей поддиапазонов, ассоциативно связанные с одним или более сигналами источников, предназначенных для повторного микширования;

определяют один или более наборов весов на основании коэффициентов усиления, оценок мощностей поддиапазонов и набора параметров микширования и

оценивают второй набор сигналов поддиапазонов с использованием, по меньшей мере, одного набора весов.

4. Способ по п.3, в котором определение одного или более наборов весов дополнительно содержит этапы, на которых:

определяют набор весов, который минимизирует разность между исходным многоканальным звуковым сигналом и повторно микшированным многоканальным звуковым сигналом.

5. Способ по п.3, в котором определение одного или более наборов весов дополнительно содержит этапы, на которых:

формируют систему линейных уравнений, в которой каждое уравнение в системе является суммой произведений, а каждое произведение сформировано перемножением сигнала поддиапазона с весом; и

определяют вес, решая систему линейных уравнений.

6. Способ по п.3, дополнительно содержащий этапы, на которых:

настраивают один или более контрольных сигналов перепадов уровней, ассоциативно связанных со вторым набором сигналов поддиапазонов, чтобы соответствовать одному или более контрольных сигналов перепадов уровней, ассоциативно связанных с первым набором сигналов поддиапазонов.

7. Способ по п.3, дополнительно содержащий этапы, на которых:

ограничивают оценку мощности поддиапазона повторно микшированного многоканального звукового сигнала, чтобы была большей или равной пороговому значению ниже оценки мощности поддиапазона исходного многоканального звукового сигнала.

8. Способ по п.3, дополнительно содержащий этапы, на которых:

масштабируют оценки мощностей поддиапазонов значением, большим, чем таковое до использования оценок мощностей поддиапазонов, чтобы определять один или более наборов весов.

9. Способ по п.1, дополнительно содержащий этапы, на которых:

модифицируют степень окружения исходного многоканального звукового сигнала с использованием оценок мощностей поддиапазонов и набора параметров микширования.

10. Способ по п.1, в котором получение набора параметров микширования дополнительно содержит этапы, на которых:

получают заданные пользователем значения усиления и панорамирования и определяют набор параметров микширования по значениям усиления и панорамирования, а также дополнительной информации.

11. Способ по п.1, в котором получение набора параметров микширования дополнительно содержит этапы, на которых:  
принимают пользовательский ввод, задающий набор параметров микширования.
12. Способ по п.3, в котором определение одного или более наборов весов  
5 дополнительно содержит этапы, на которых:  
определяют амплитуду первого набора весов и  
определяют амплитуду второго набора весов, при этом второй набор весов  
включает в себя иное количество весов, чем первый набор весов.
13. Способ по п.12, дополнительно содержащий этапы, на которых:  
10 сравнивают амплитуды первого и второго наборов весов и  
выбирают один из первого и второго наборов весов для использования при  
оценивании второго набора сигналов поддиапазонов на основании результатов  
сравнения.
14. Способ по п.3, который дополнительно содержит этапы, на которых  
15 сглаживают один или более наборов весов по времени.
15. Способ по п.14, который дополнительно содержит этапы, на которых  
регулируют сглаживание одного или более наборов весов по времени, чтобы  
20 уменьшать звуковые искажения.
16. Способ по п.14, который дополнительно содержит этапы, на которых:  
определяют, превышает ли тональный или стационарный показатель исходного  
многоканального звукового сигнала пороговое значение; и  
25 сглаживают один или более наборов весов по времени, если показатель превышает  
пороговое значение
17. Способ кодирования многоканального звукового сигнала, предназначенного  
для повторного микширования, содержащий этапы, на которых:  
получают звуковой сигнал, содержащий набор объектов;  
30 получают сигналы источников, включающие в себя один или более сигналов  
источников, предназначенных для повторного микширования; и  
формируют дополнительную информацию из сигналов источников и звукового  
сигнала, представляющую, по меньшей мере, одно из зависимости между звуковым  
сигналом и каждым из сигналов источников и энергетического уровня каждого из  
35 сигналов источников.
18. Способ по п.17, в котором формирование дополнительной информации  
дополнительно содержит этапы, на которых:  
раскладывают звуковой сигнал и подмножество сигналов источников на первый  
40 набор сигналов поддиапазонов и второй набор сигналов поддиапазонов,  
соответственно;  
причем для каждого сигнала поддиапазона во втором наборе сигналов  
поддиапазонов:  
оценивают мощность поддиапазона для сигнала поддиапазона;  
45 получают один или более коэффициентов усиления и  
формируют дополнительную информацию из одного или более коэффициентов  
усиления и мощности поддиапазона.
19. Способ декодирования многоканального звукового сигнала с повторным  
50 микшированием, содержащий этапы, на которых:  
получают многоканальный звуковой сигнал;  
определяют коэффициенты усиления для набора сигналов источников с  
использованием требуемых перепадов уровней источников, представляющих

требуемые направления звучания набора сигналов источников в павильоне звукозаписи;

оценивают мощность поддиапазона для направления прямого звучания набора сигналов источников с использованием многоканального звукового сигнала и

оценивают мощности поддиапазонов для, по меньшей мере, некоторых из сигналов источников в наборе сигналов источников, модифицируя мощность поддиапазона для направления прямого звучания в качестве функции направления прямого звучания и требуемого направления звучания.

20. Способ по п.19, в котором функция является функцией направления звучания, которая возвращает коэффициент усиления около единицы только для требуемого направления звучания.

21. Способ декодирования многоканального звукового сигнала с повторным микшированием, содержащий этапы, на которых:

получают микшированный звуковой сигнал;

получают набор параметров микширования для повторного микширования микшированного звукового сигнала;

если дополнительная информация доступна,

повторно микшируют микшированный звуковой сигнал с использованием дополнительной информации и набора параметров микширования;

если дополнительная информация недоступна,

формируют набор слепых параметров из микшированного звукового сигнала и

формируют повторно микшированный звуковой сигнал с использованием слепых параметров и набора параметров микширования.

22. Способ по п.21, дополнительно содержащий этапы, на которых:

формируют параметры повторного микширования либо из слепых параметров, либо из дополнительной информации и, если параметры повторного микширования

формируются из дополнительной информации,

формируют повторно микшированный звуковой сигнал из параметров повторного микширования и микшированного сигнала.

23. Способ по п.21, дополнительно содержащий этапы, на которых

добавляют один или более эффектов в повторно микшированный звуковой сигнал.

24. Способ декодирования многоканального звукового сигнала с повторным микшированием, содержащий этапы, на которых:

получают микшированный звуковой сигнал, включающий в себя сигналы речевого источника;

получают параметры микширования, задающие улучшенное усовершенствование одному или более речевым сигналам источников;

формируют набор слепых параметров из микшированного звукового сигнала;

формируют параметры повторного микширования из слепых параметров и параметров микширования и

применяют параметры повторного микширования к микшированному сигналу, чтобы усовершенствовать один или более речевых сигналов источников в соответствии с параметрами микширования.

25. Способ декодирования многоканального звукового сигнала с повторным микшированием, содержащий этапы, на которых:

формируют пользовательский интерфейс для приема входных данных, задающих параметры микширования;

получают параметр микширования через пользовательский интерфейс;

получают исходный звуковой сигнал, включающий в себя сигналы источников; получают дополнительную информацию, представляющую зависимость между исходным звуковым сигналом и одним или более сигналами источников; и повторно микшируют один или более сигналов источников с использованием

5 дополнительной информации и параметра микширования, чтобы сформировать повторно микшированный звуковой сигнал.

26. Способ по п.25, дополнительно содержащий этапы, на которых принимают исходный звуковой сигнал или дополнительную информацию из

10 сетевого ресурса.

27. Способ декодирования многоканального звукового сигнала с повторным микшированием, содержащий этапы, на которых:

получают микшированный звуковой сигнал;

получают набор параметров микширования для повторного микширования

15 микшированного звукового сигнала;

формируют параметры повторного микширования с использованием микшированного звукового сигнала и набора параметров микширования и формируют повторно микшированный звуковой сигнал, применяя параметры

20 повторного микширования к микшированному звуковому сигналу с использованием матрицы nхn.

28. Устройство для декодирования многоканального звукового сигнала с повторным микшированием, содержащее:

декодер, сконфигурированный для приема исходного многоканального звукового

25 сигнала, содержащего набор сигналов источников, который включает в себя один или более сигналов источников, предназначенных для повторного микширования, и для приема дополнительной информации, представляющей, по меньшей мере, одно из зависимости между исходным многоканальным звуковым сигналом и каждым из

30 набора сигналов источников, и энергетического уровня каждого из набора сигналов источников;

интерфейс, сконфигурированный для получения набора параметров микширования для повторного микширования исходного многоканального звукового сигнала; и

модуль повторного микширования, соединенный с декодером и интерфейсом,

35 причем модуль повторного микширования является конфигурируемым для повторного микширования сигналов источников с использованием дополнительной информации и набора параметров микширования, чтобы сформировать повторно микшированный многоканальный звуковой сигнал.

29. Устройство по п.28, в котором набор параметров микширования задается

40 пользователем через интерфейс.

30. Устройство по п.29, в котором модуль повторного микширования оценивает второй набор сигналов поддиапазонов, соответствующий повторно микшированному многоканальному звуковому сигналу, с использованием дополнительной информации

45 и набора параметров микширования и преобразует второй набор сигналов поддиапазонов в повторно микшированный многоканальный звуковой сигнал.

31. Устройство по п.29, в котором декодер декодирует дополнительную информацию для выдачи коэффициентов усиления и оценок мощностей

50 поддиапазонов, ассоциативно связанных с сигналами источников, предназначенных для повторного микширования, а модуль повторного микширования определяет один или более наборов весов на основании коэффициентов усиления, оценок мощности поддиапазонов и набора параметров микширования и оценивает второй набор

сигналов поддиапазонов с использованием по меньшей мере одного набора весов.

32. Устройство по п.31, в котором модуль повторного микширования определяет один или более наборов весов, определяя амплитуду первого набора весов и определяя амплитуду второго набора весов, причем второй набор весов включает в себя иное количество весов, чем первый набор весов.

33. Устройство по п.32, в котором модуль повторного микширования сравнивает амплитуды первого и второго наборов весов и выбирает один из первого и второго наборов весов для использования при оценивании второго набора сигналов поддиапазонов на основании результатов сравнения.

34. Устройство по п.31, в котором модуль повторного микширования определяет один или более наборов весов, определяя набор весов, который минимизирует разность между исходным многоканальным звуковым сигналом и повторно микшированным многоканальным звуковым сигналом.

35. Устройство по п.31, в котором модуль повторного микширования настраивает один или более контрольных сигналов перепадов уровней, ассоциативно связанных со вторым набором сигналов поддиапазонов, чтобы соответствовать одному или более контрольным сигналам перепадов уровней, ассоциативно связанным с первым набором сигналов поддиапазонов.

36. Устройство по п.31, в котором модуль повторного микширования ограничивает оценку мощности поддиапазона повторно микшированного многоканального звукового сигнала, чтобы была большей или равной пороговому значению ниже оценки мощности поддиапазона исходного многоканального звукового сигнала.

37. Устройство по п.31, в котором модуль повторного микширования масштабирует оценки мощностей поддиапазонов значением, большим, чем таковое до использования оценок мощностей поддиапазонов, чтобы определять один или более наборов весов.

38. Устройство по п.31, в котором модуль повторного микширования сглаживает один или более наборов весов по времени.

39. Устройство по п.38, в котором модуль повторного микширования регулирует сглаживание одного или более наборов весов по времени, чтобы уменьшать звуковые искажения.

40. Устройство по п.38, в котором модуль повторного микширования сглаживает один или более наборов весов по времени на основании тонального или стационарного показателя.

41. Устройство по п.38, в котором модуль повторного микширования определяет, превышает ли тональный или стационарный показатель исходного многоканального звукового сигнала пороговое значение; и сглаживает один или более наборов весов по времени, если показатель превышает пороговое значение.

42. Устройство кодирования многоканального звукового сигнала, предназначенного для повторного микширования, содержащее:

интерфейс, конфигурируемый для получения звукового сигнала, содержащего набор объектов и сигналы источников, представляющие объекты;

и

формирователь дополнительной информации, соединенный с интерфейсом и конфигурируемый для формирования дополнительной информации из сигналов источников, причем дополнительная информация представляет взаимосвязь между звуковым сигналом и сигналами источников.

43. Устройство кодирования многоканального звукового сигнала,

предназначенного для повторного микширования, содержащее:

интерфейс, конфигурируемый для получения многоканального звукового сигнала; и  
формирователь дополнительной информации, конфигурируемый для определения  
5 коэффициентов усиления для набора сигналов источников с использованием  
требуемых перепадов уровней источников, представляющих требуемые направления  
звучания набора сигналов источников в павильоне звукозаписи, оценивания  
10 мощностей поддиапазона для направления прямого звучания набора сигналов  
источников с использованием многоканального звукового сигнала и оценивания  
мощностей поддиапазонов для, по меньшей мере, некоторых из сигналов источников  
в наборе сигналов источников посредством модифицирования мощности  
поддиапазона для направления прямого звучания в качестве функции направления  
прямого звучания и требуемого направления звучания.

44. Устройство для декодирования многоканального звукового сигнала с  
15 повторным микшированием, содержащее:

формирователь параметров, конфигурируемый для получения микшированного  
звукового сигнала и набора параметров микширования для повторного  
микширования звукового сигнала и для определения того, доступна ли  
20 дополнительная информация; и

формирователь повторно микшированного сигнала, соединенный с  
формирователем параметров и конфигурируемый для повторного микширования  
микшированного звукового сигнала с использованием дополнительной информации и  
набора параметров микширования, если дополнительная информация доступна, а если  
25 дополнительная информация недоступна, приема набора слепых параметров и  
формирования повторно микшированного звукового сигнала с использованием  
слепых параметров и набора параметров микширования.

45. Устройство для декодирования многоканального звукового сигнала с  
30 повторным микшированием, содержащее:

интерфейс, конфигурируемый для получения микшированного звукового сигнала,  
включающего в себя сигналы речевых источников и параметры микширования,  
задающие улучшенное усовершенствование для одного или более речевых сигналов  
источников;

формирователь параметров микширования, соединенный с интерфейсом и  
конфигурируемый для формирования набора слепых параметров из микшированного  
звукового сигнала и для формирования параметров из слепых параметров и  
параметров микширования; и

формирователь повторно микшированного сигнала, конфигурируемый для  
40 применения параметров к микшированному сигналу, чтобы усовершенствовать один  
или более речевых сигналов источников в соответствии с параметрами микширования.

46. Устройство для декодирования многоканального звукового сигнала с  
повторным микшированием, содержащее:

интерфейс, конфигурируемый для получения набора параметров микширования для  
45 повторного микширования микшированного звукового сигнала; и

модуль повторного микширования, соединенный с интерфейсом и  
конфигурируемый для формирования параметров повторного микширования с  
использованием микшированного звукового сигнала и набора параметров  
50 микширования и для формирования повторно микшированного звукового сигнала  
посредством применения параметров повторного микширования к микшированному  
звуковому сигналу с использованием матрицы  $n \times n$ .

47. Машиночитаемый носитель, содержащий команды, хранимые на нем, которые, когда выполняются процессором, предписывают процессору выполнять операции, содержащие:

5 получение исходного многоканального звукового сигнала, содержащего набор сигналов источников, который включает в себя один или более сигналов источников, предназначенных для повторного микширования;

10 получение дополнительной информации, представляющей, по меньшей мере, одно из зависимости между исходным многоканальным звуковым сигналом и каждым из набора сигналов источников, и энергетического уровня каждого из набора сигналов источников;

получение набора параметров микширования для повторного микширования исходного многоканального звукового сигнала и

15 формирование повторно микшированного многоканального звукового сигнала с использованием исходного многоканального звукового сигнала, дополнительной информации и набора параметров микширования.

48. Машиночитаемый носитель, содержащий команды, хранимые на нем, которые, когда выполняются процессором, предписывают процессору выполнять операции, содержащие:

20 получение звукового сигнала, содержащего набор объектов;

получение сигналов источников, включающих в себя один или более сигналов источников, предназначенных для повторного микширования; и

25 формирование дополнительной информации из сигналов источников и звукового сигнала, представляющей по меньшей мере одно из зависимости между звуковым сигналом и каждым из сигналов источников, и энергетического уровня каждого из набора сигналов источников.

49. Система для декодирования многоканального звукового сигнала с повторным микшированием, содержащая:

30 процессор и

машиночитаемый носитель, соединенный с процессором и включающий в себя команды, которые, когда выполняются процессором, предписывают процессору выполнять операции, содержащие:

35 получение исходного многоканального звукового сигнала, содержащего набор сигналов источников, который включает в себя один или более сигналов источников, предназначенных для повторного микширования;

40 получение дополнительной информации, представляющей, по меньшей мере, одно из зависимости между исходным многоканальным звуковым сигналом и каждым из набора сигналов источников, и энергетического уровня каждого из набора сигналов источников;

получение набора параметров микширования для повторного микширования исходного многоканального звукового сигнала и

45 формирование повторно микшированного многоканального звукового сигнала с использованием исходного многоканального звукового сигнала, дополнительной информации и набора параметров микширования.

50 Система для декодирования многоканального звукового сигнала с повторным микшированием, содержащая:

процессор и

машиночитаемый носитель, присоединенный к процессору и включающий в себя команды, которые, когда выполняются процессором, предписывают процессору

выполнять операции, содержащие:

получение звукового сигнала, содержащего набор объектов;

получение сигналов источников, включающих в себя один или более сигналов источников, предназначенных для повторного микширования;

5

и

формирование дополнительной информации из сигналов источников и звукового сигнала, представляющей, по меньшей мере, одно из зависимости между звуковым сигналом и каждым из сигналов источников, и энергетического уровня каждого из

10

набора сигналов источников.

15

20

25

30

35

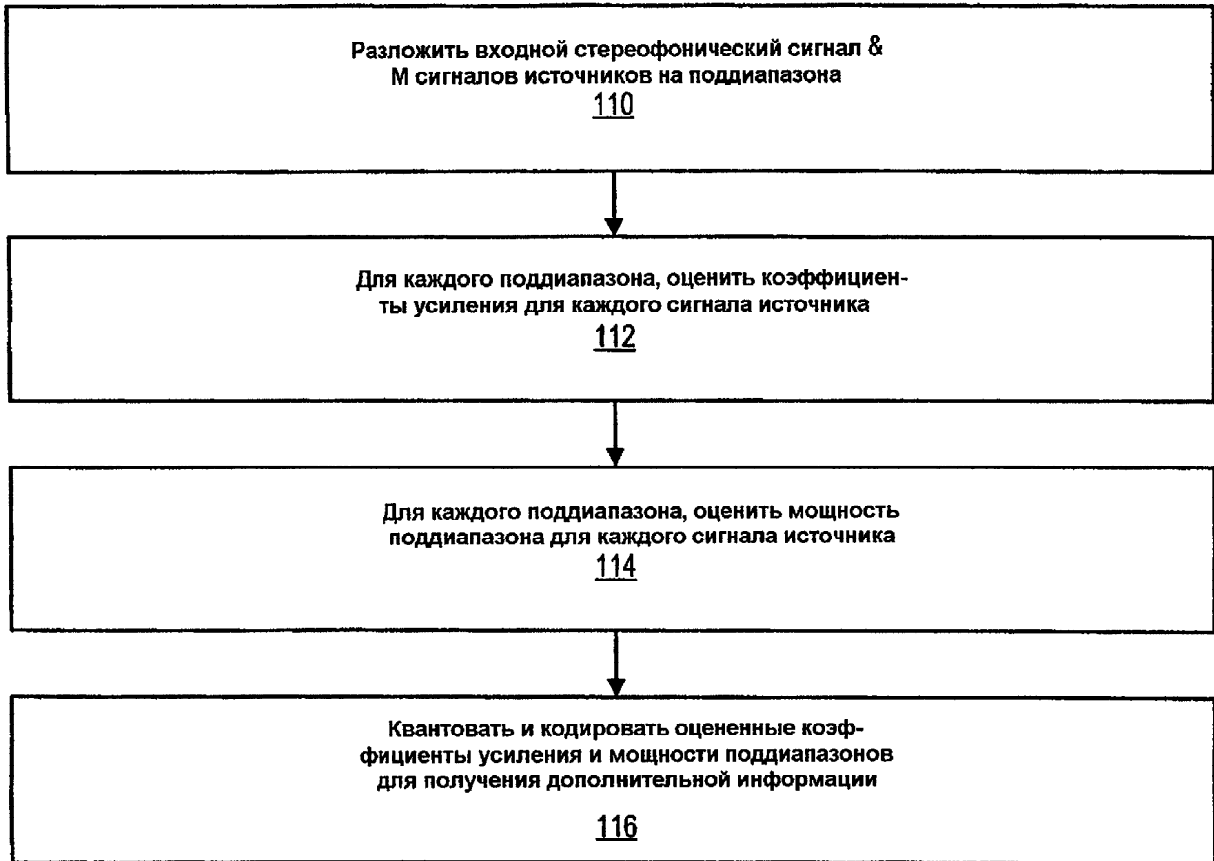
40

45

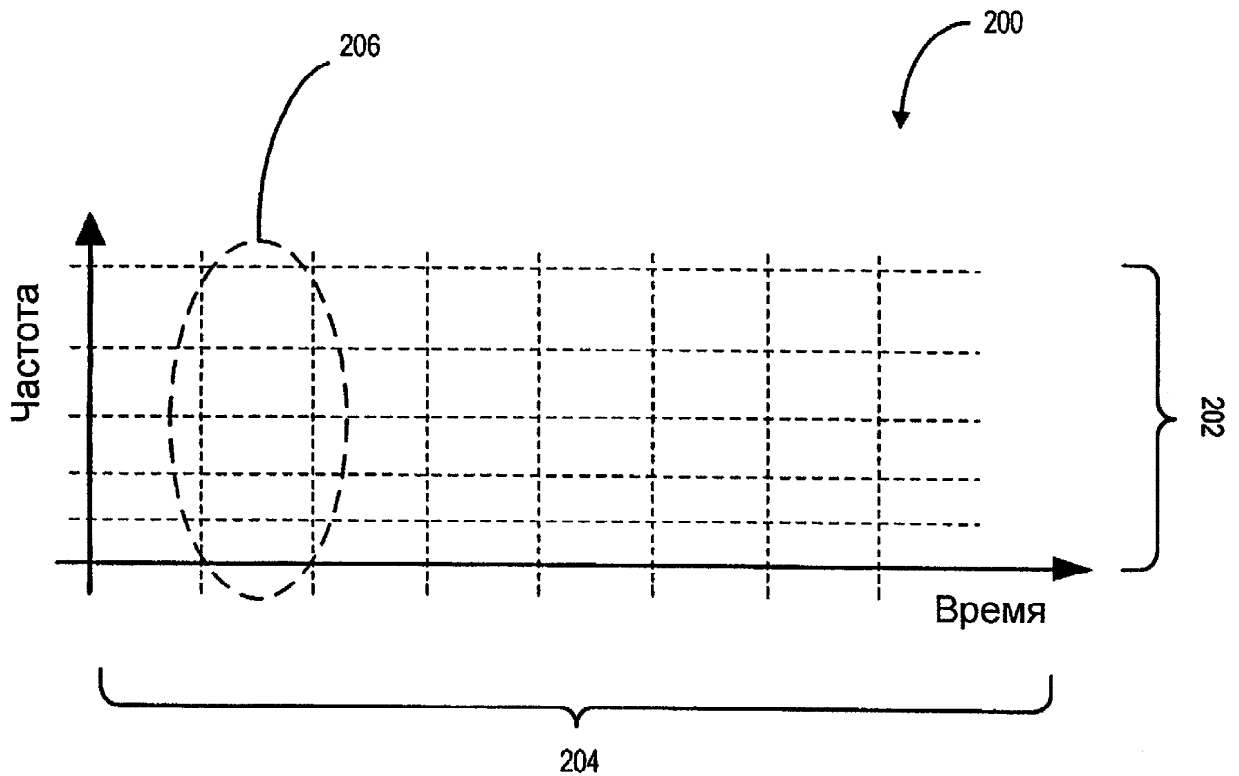
50

Последовательность операций формирования  
дополнительной информации (доступны M  
сигналов источников)

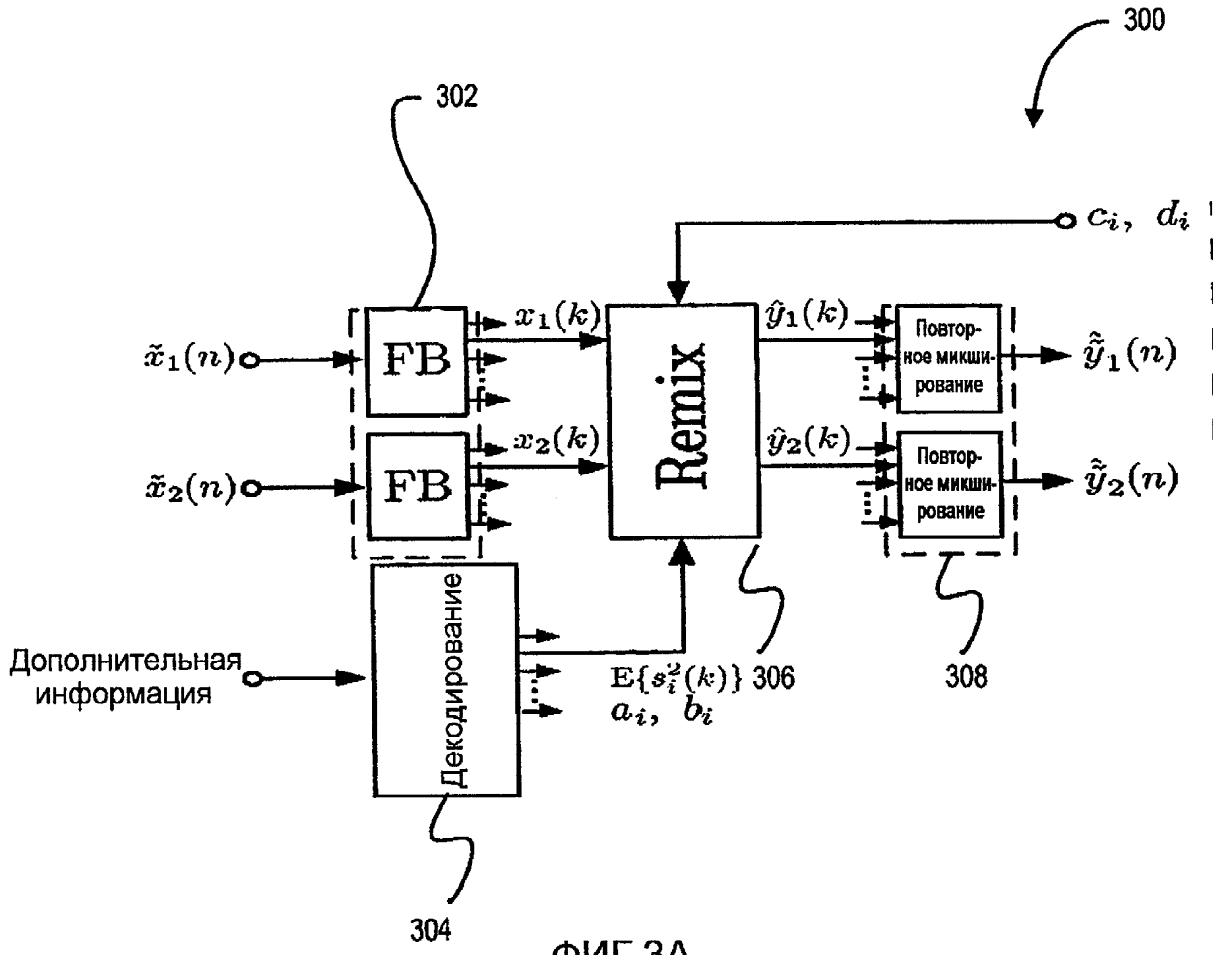
108



ФИГ.1В

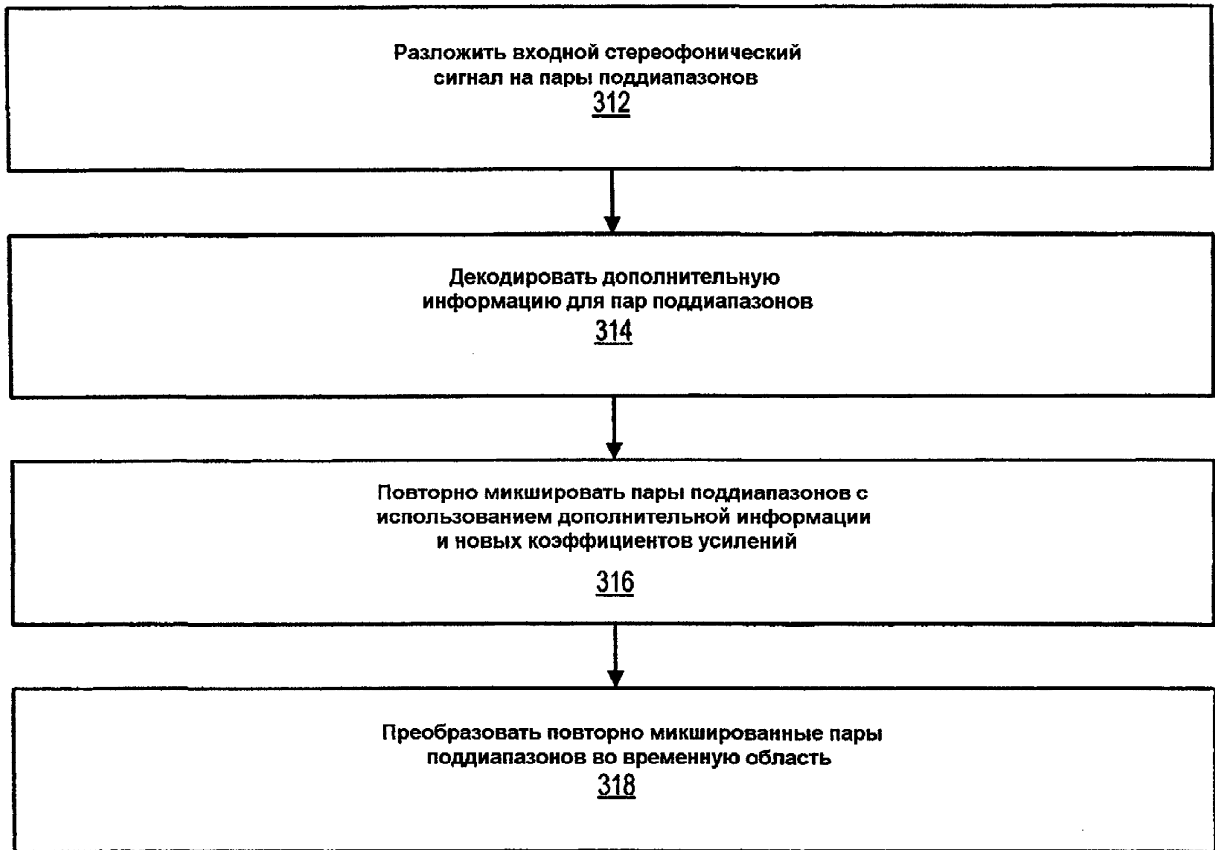


ФИГ.2

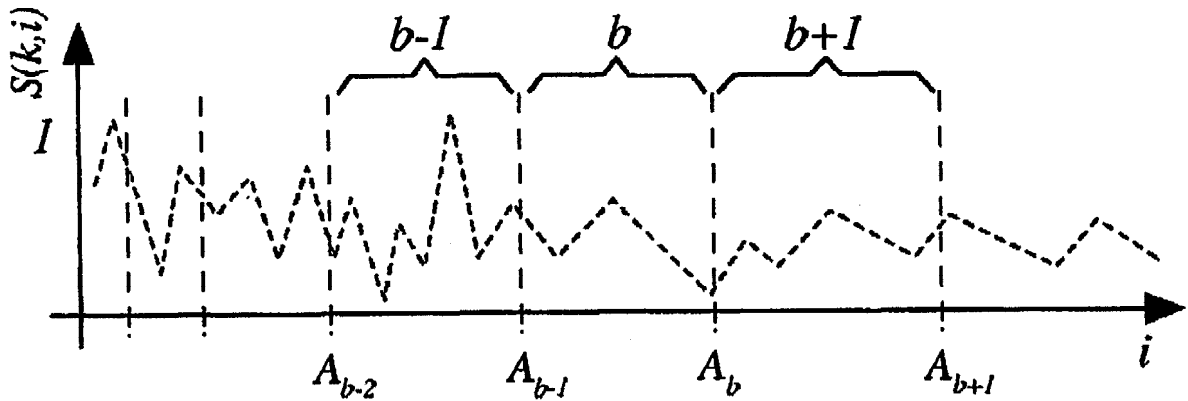


ФИГ.3А

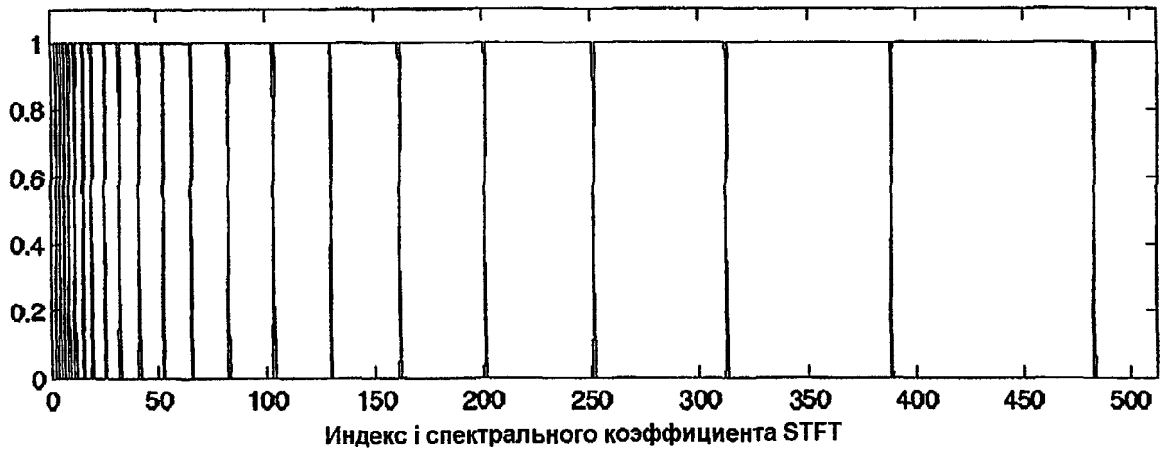
Последовательность операций  
повторного микширования  
310



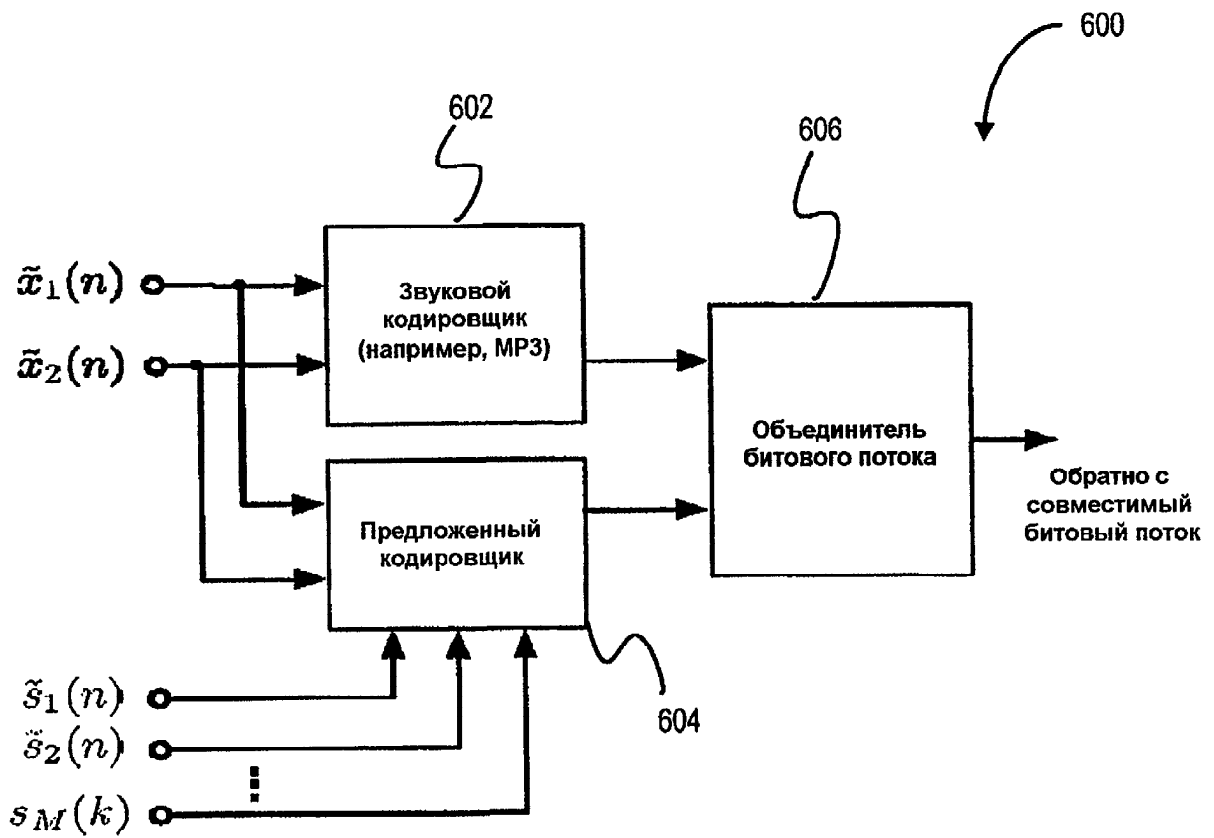
ФИГ.3В



ФИГ.4

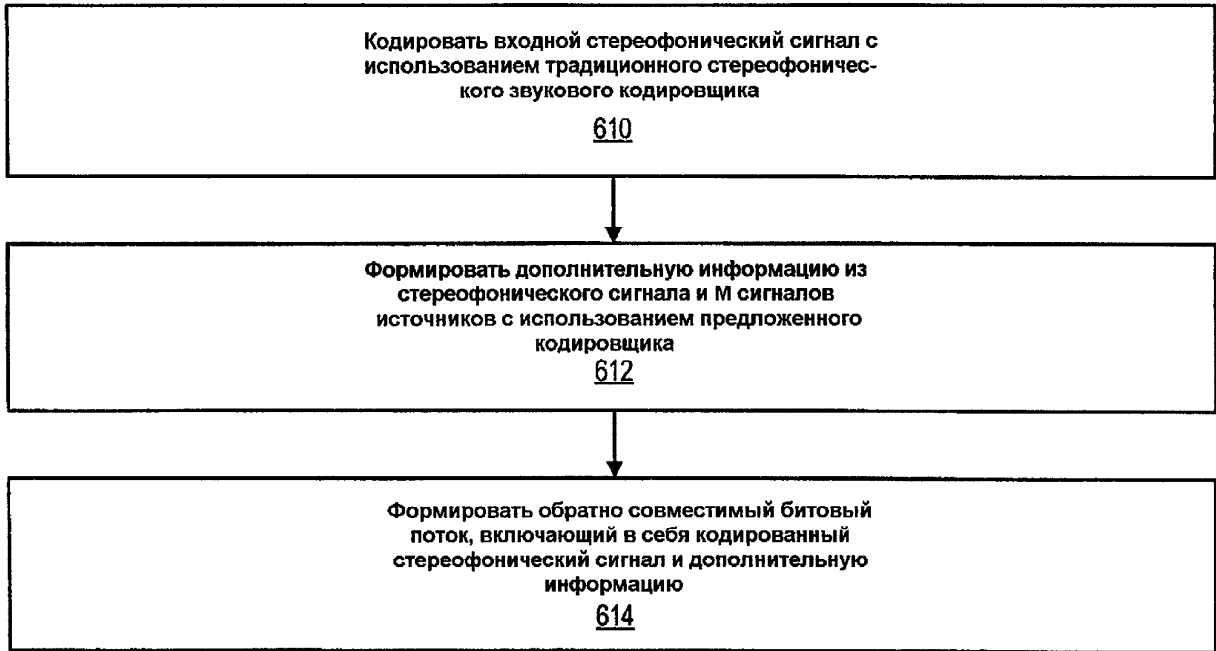


ФИГ.5

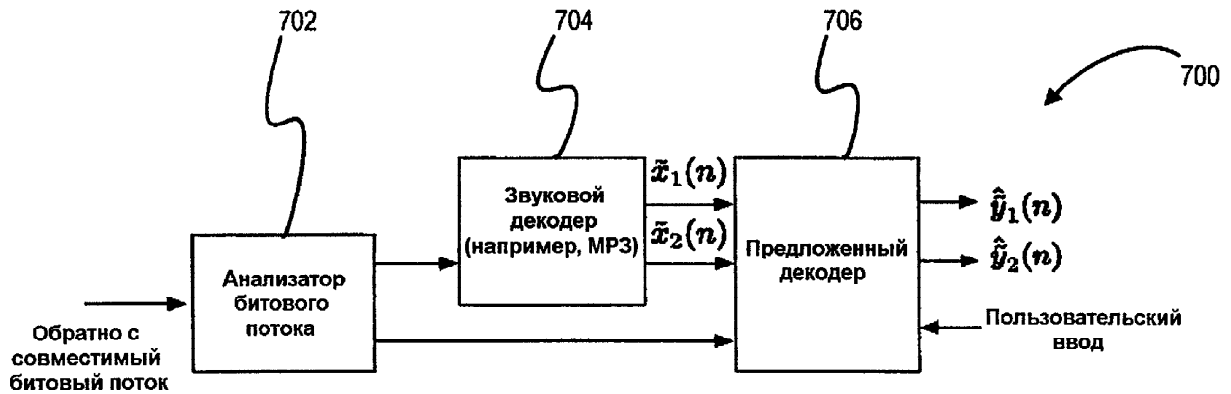


ФИГ.6А

Последовательность операций формирования дополнительной информации (объединенные предложенный кодировщик и традиционный звуковой кодировщик)  
608



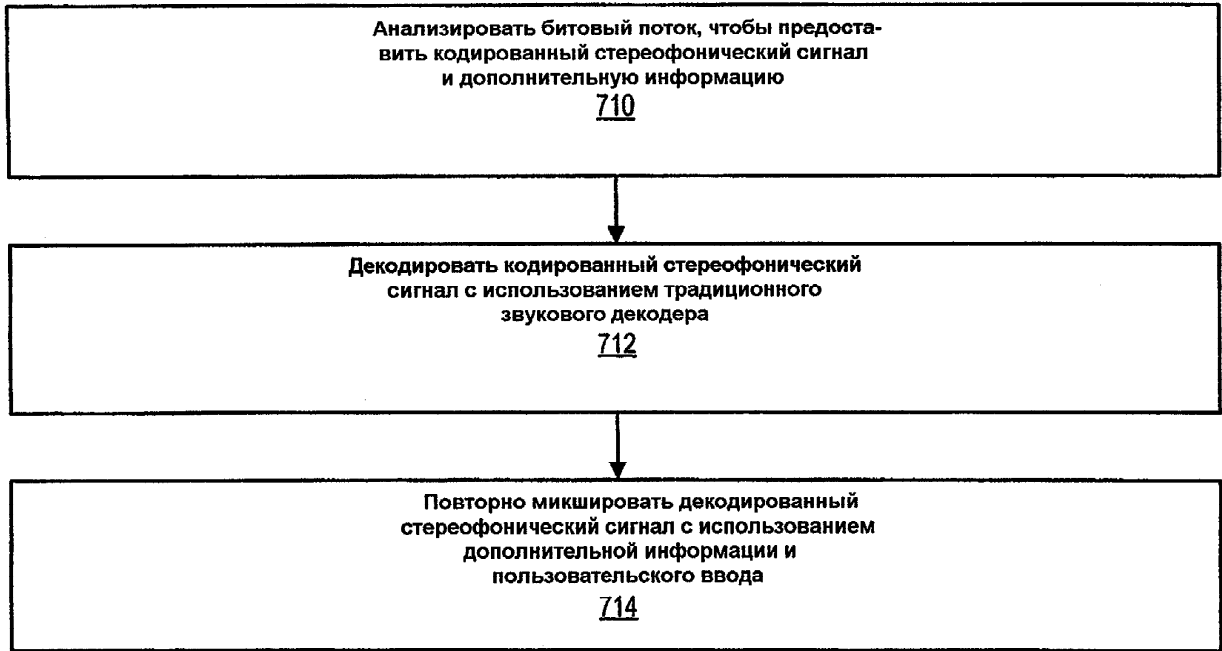
ФИГ.6В



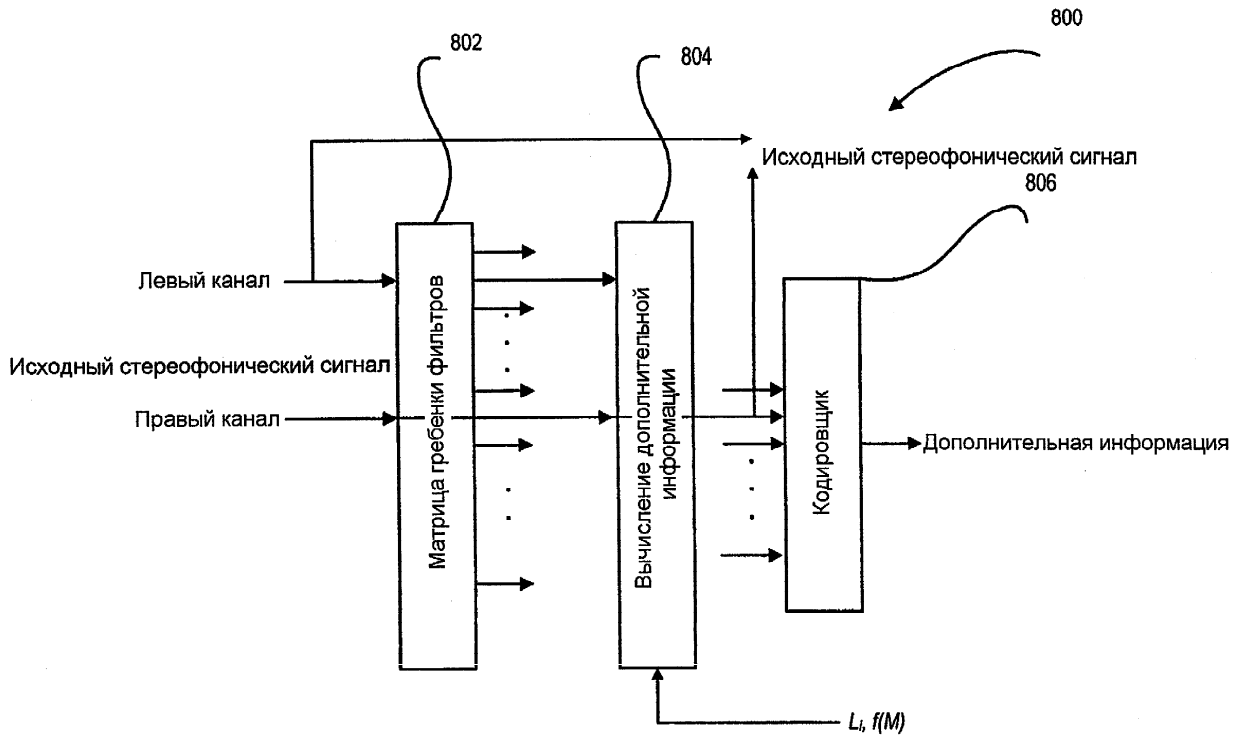
ФИГ.7А

Последовательность операций повторного микширования (объединенный предложенный декодер и традиционный звуковой декодер)

708



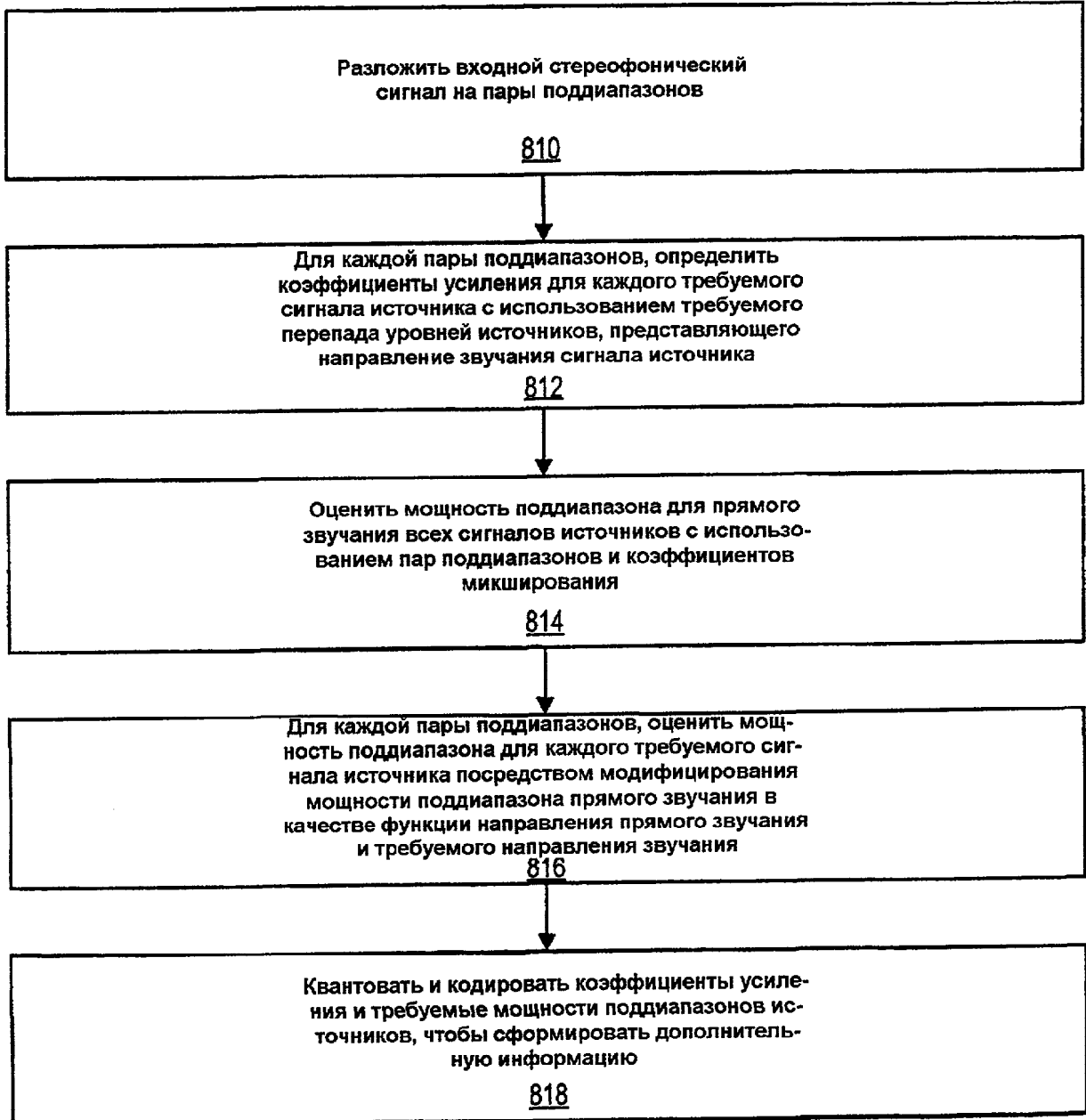
ФИГ.7В



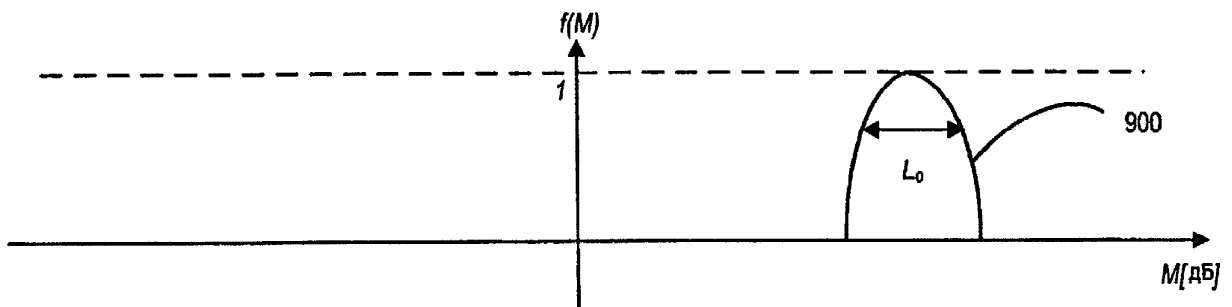
ФИГ.8А

Последовательность операций формирования  
дополнительной информации

808

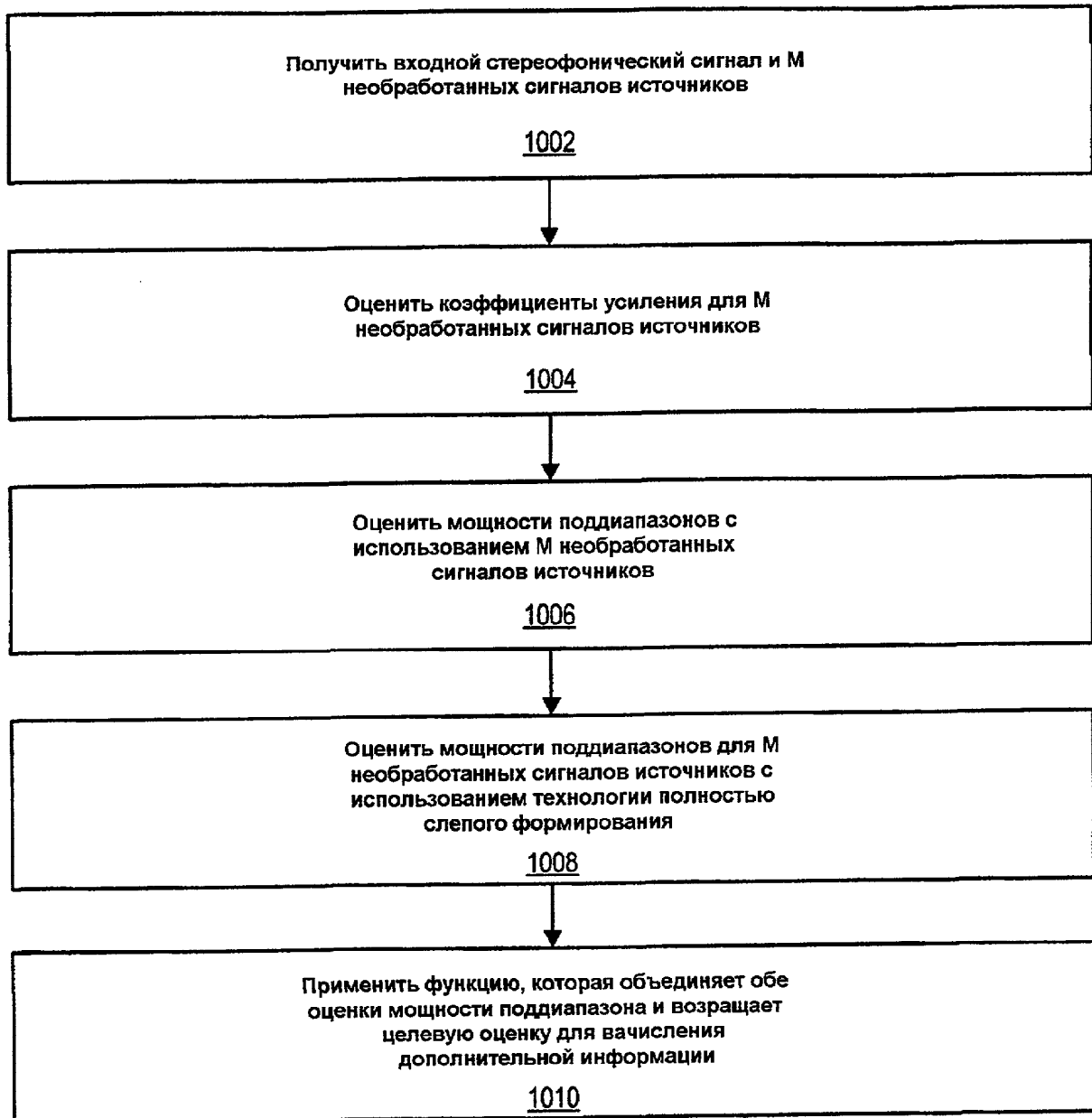


ФИГ.8В

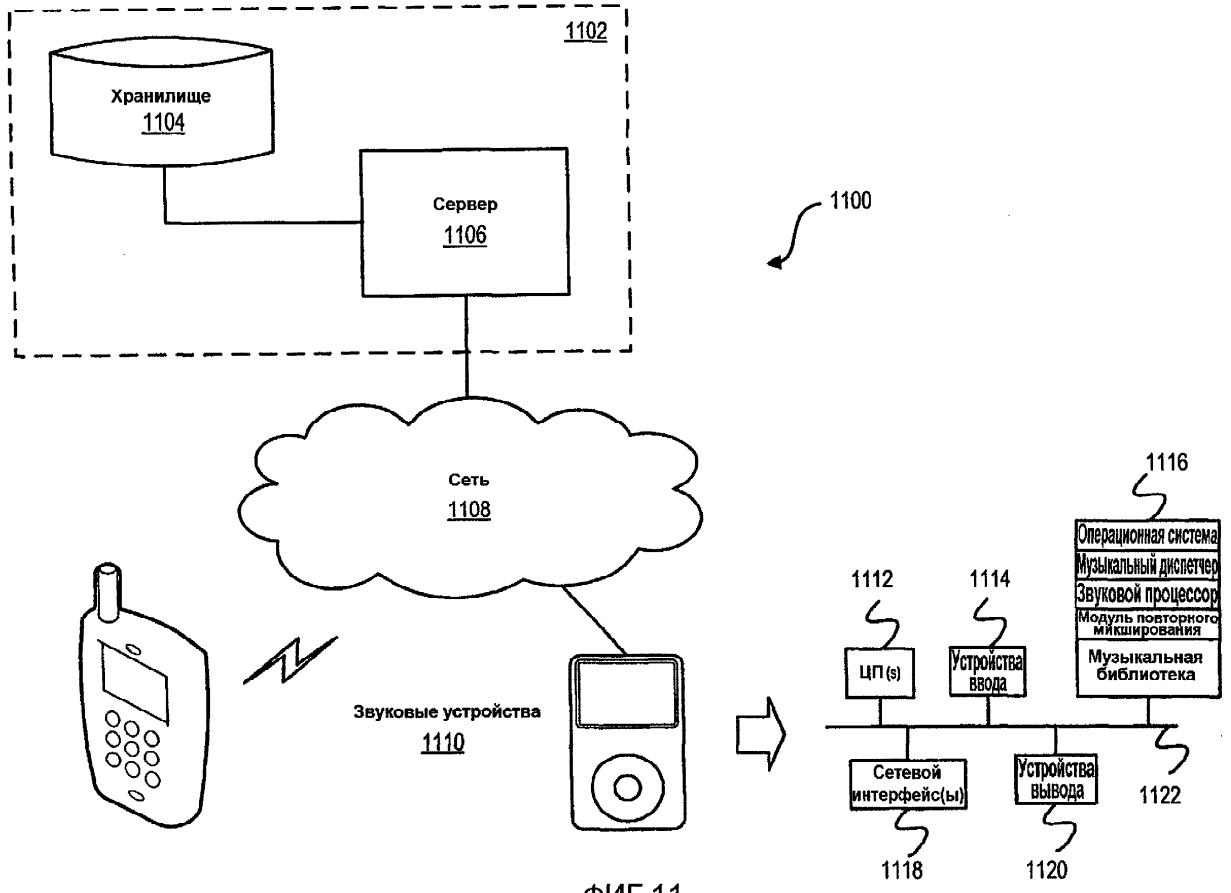


ФИГ.9

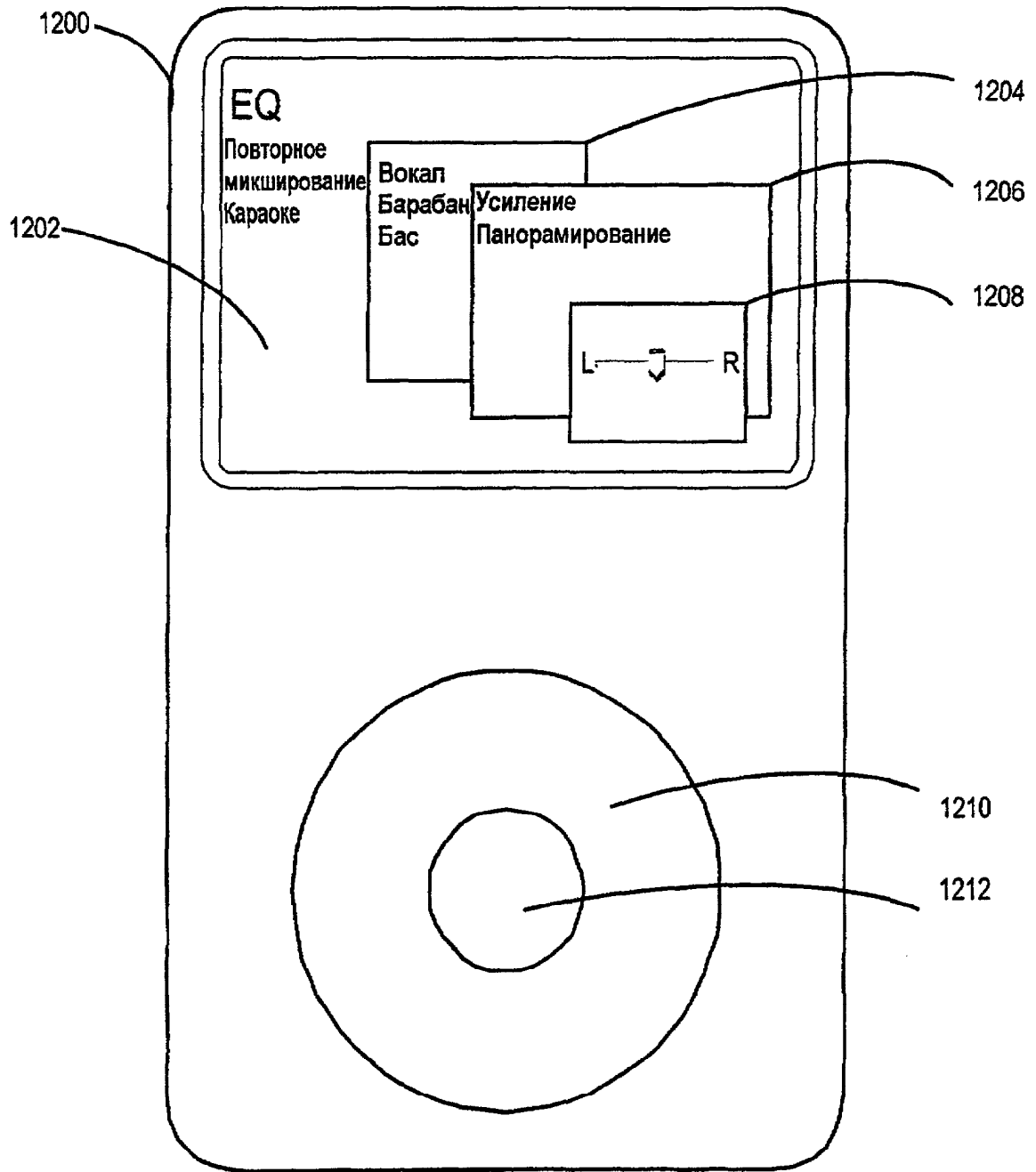
Последовательность операций формирования  
дополнительной информации (частично  
слепое формирование)  
1000



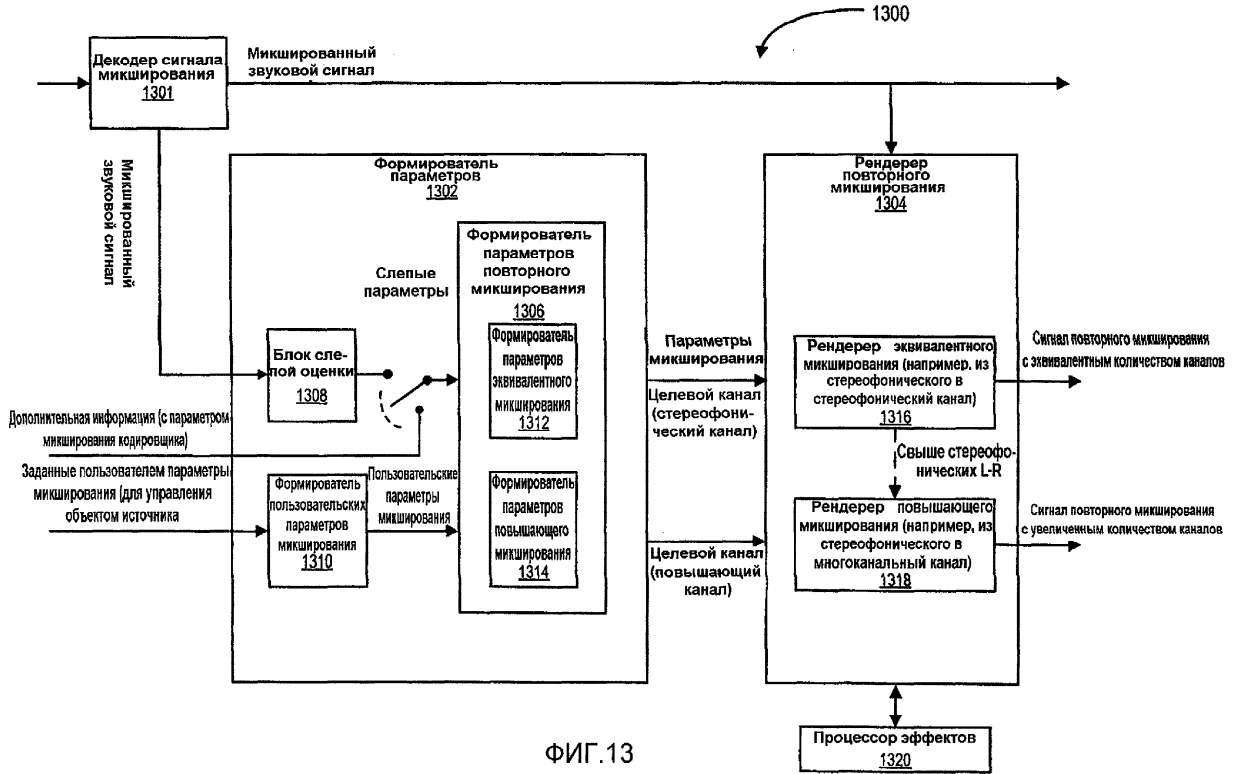
ФИГ.10



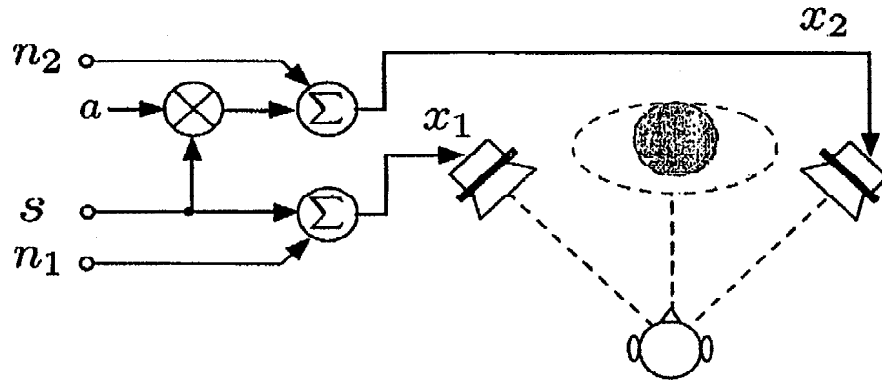
ФИГ.11



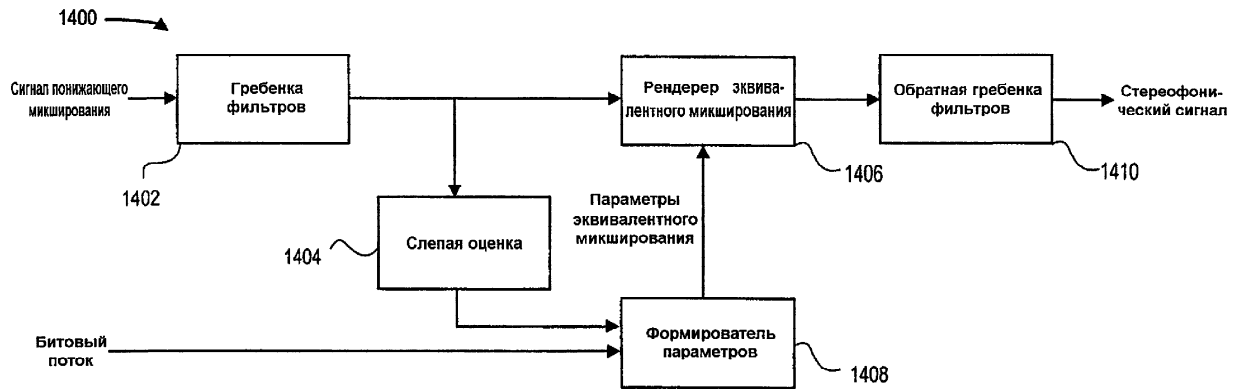
ФИГ.12



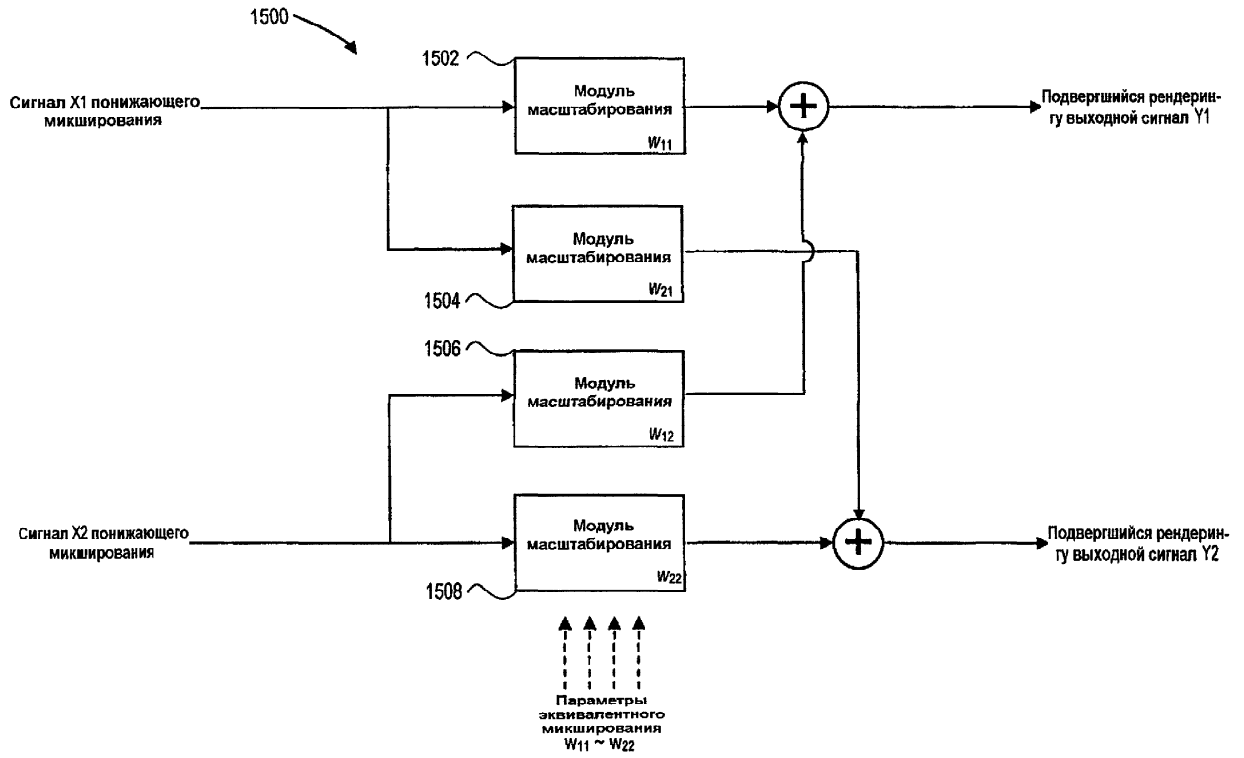
ФИГ.13



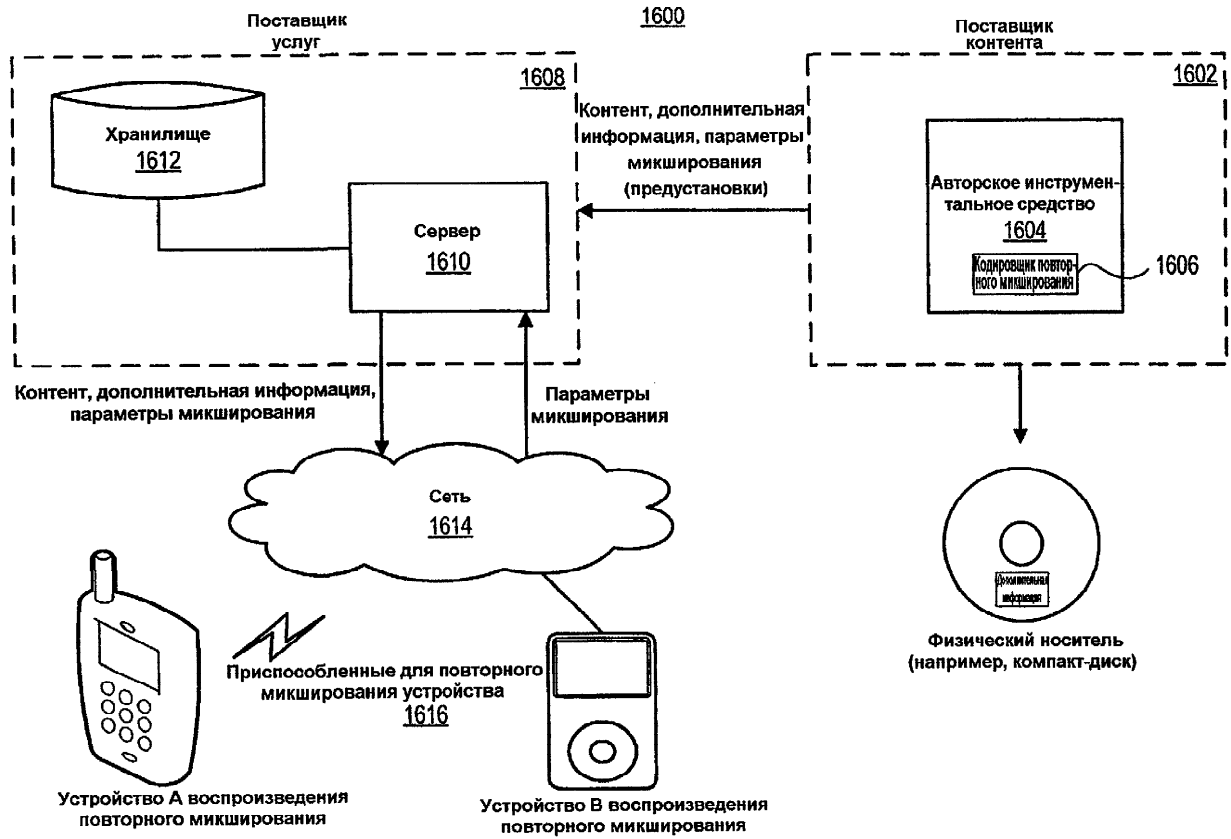
ФИГ.14А



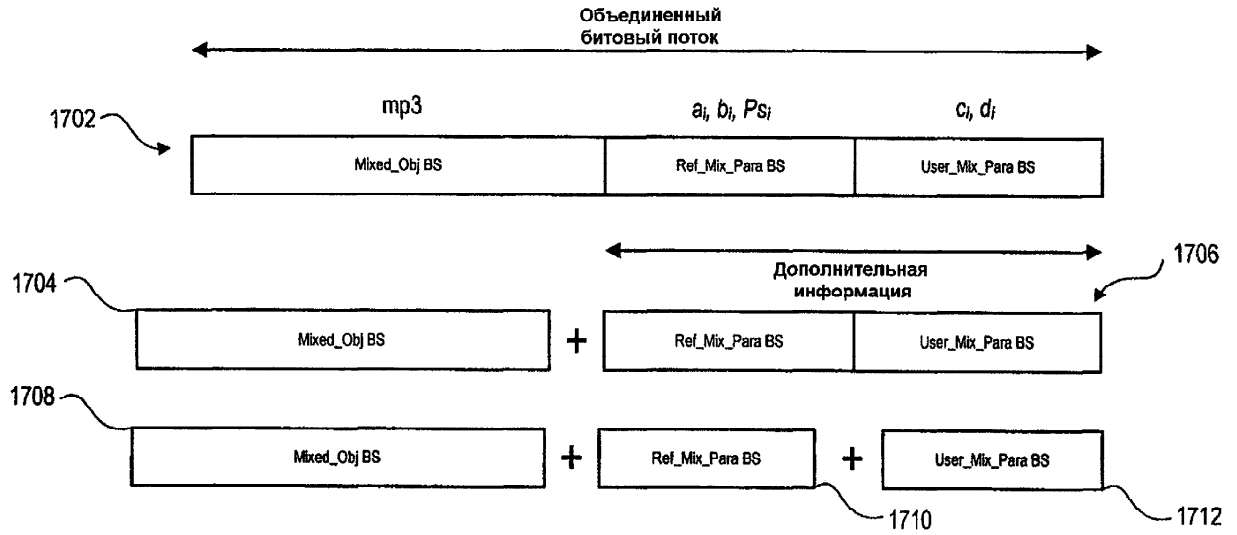
ФИГ.14В



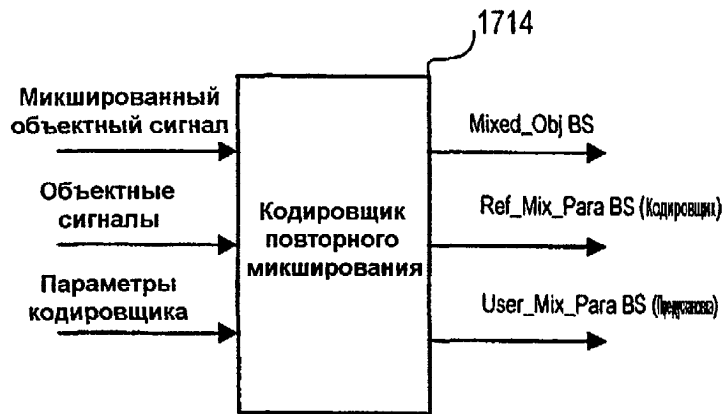
ФИГ. 15



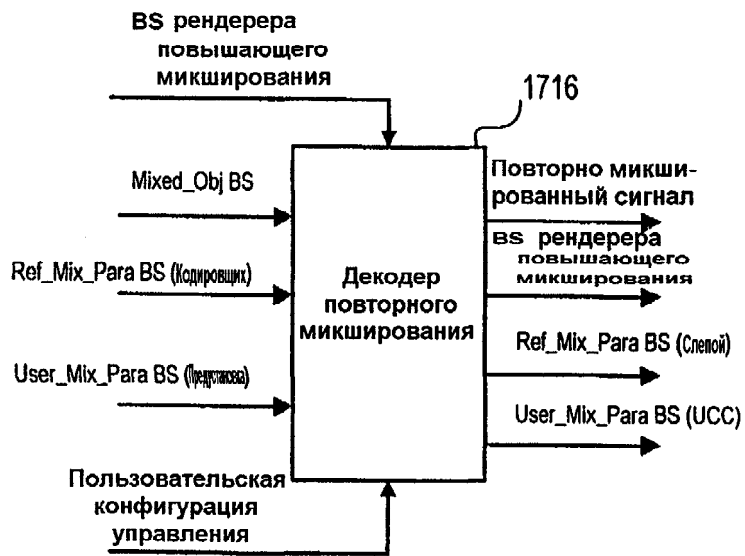
ФИГ. 16



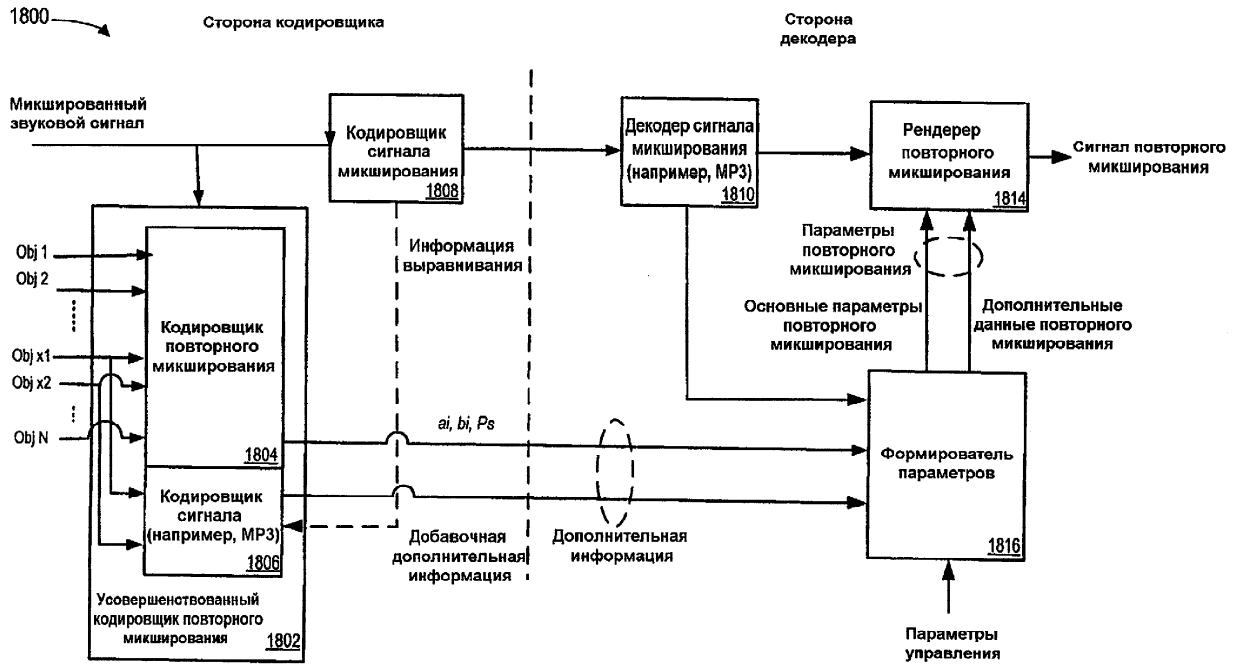
ФИГ.17А



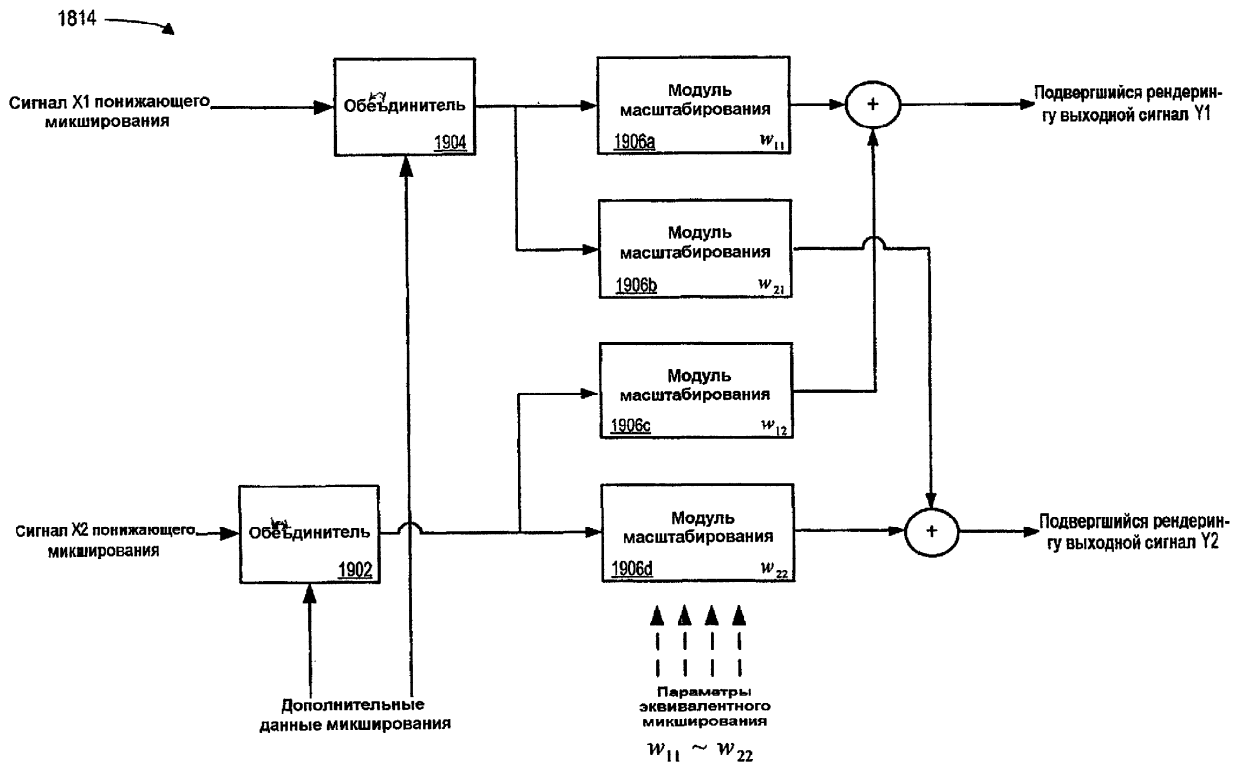
ФИГ.17В



ФИГ.17С



ФИГ.18



ФИГ.19