



ФЕДЕРАЛЬНАЯ СЛУЖБА
ПО ИНТЕЛЛЕКТУАЛЬНОЙ СОБСТВЕННОСТИ

(12) ОПИСАНИЕ ИЗОБРЕТЕНИЯ К ПАТЕНТУ

(52) СПК

G10L 19/002 (2020.02); G10L 19/008 (2020.02); G10L 19/06 (2020.02); G10L 19/24 (2020.02)

(21)(22) Заявка: 2018114899, 22.09.2016

(24) Дата начала отсчета срока действия патента:
22.09.2016

Дата регистрации:
11.08.2020

Приоритет(ы):

(30) Конвенционный приоритет:
25.09.2015 US 62/232,589;
14.07.2016 US 62/362,360

(43) Дата публикации заявки: 25.10.2019 Бюл. № 30

(45) Опубликовано: 11.08.2020 Бюл. № 23

(85) Дата начала рассмотрения заявки РСТ на
национальной фазе: 25.04.2018

(86) Заявка РСТ:
CA 2016/051107 (22.09.2016)

(87) Публикация заявки РСТ:
WO 2017/049398 (30.03.2017)

Адрес для переписки:
129090, Москва, ул. Б. Спасская, 25, стр. 3, ООО
"Юридическая фирма Городисский и
Партнеры"

(72) Автор(ы):

ВАЙАНКУР, Томми (СА),
ЕЛИНЕК, Милан (СА)

(73) Патентообладатель(и):

ВОЙСЭЙДЖ КОРПОРЕЙШН (СА)

(56) Список документов, цитированных в отчете
о поиске: US 8103005 B2, 24.01.2012. EP 2405424
B1, 12.11.2014. US 7848932 B2, 07.12.2010. US
2014/0112482 A1, 24.04.2014. US 2013/0262130
A1, 03.10.2013. RU 2520402 C2, 27.06.2014.
MORIYA et al, "Extended Linear Prediction
Tools for Lossless Audio Coding", Proceedings
of the IEEE International Conference on
Acoustics, Speech and Signal Processing, (см.
прод.)

(54) СПОСОБ И СИСТЕМА ДЛЯ КОДИРОВАНИЯ СТЕРЕОФОНИЧЕСКОГО ЗВУКОВОГО СИГНАЛА С ИСПОЛЬЗОВАНИЕМ ПАРАМЕТРОВ КОДИРОВАНИЯ ПЕРВИЧНОГО КАНАЛА ДЛЯ КОДИРОВАНИЯ ВТОРИЧНОГО КАНАЛА

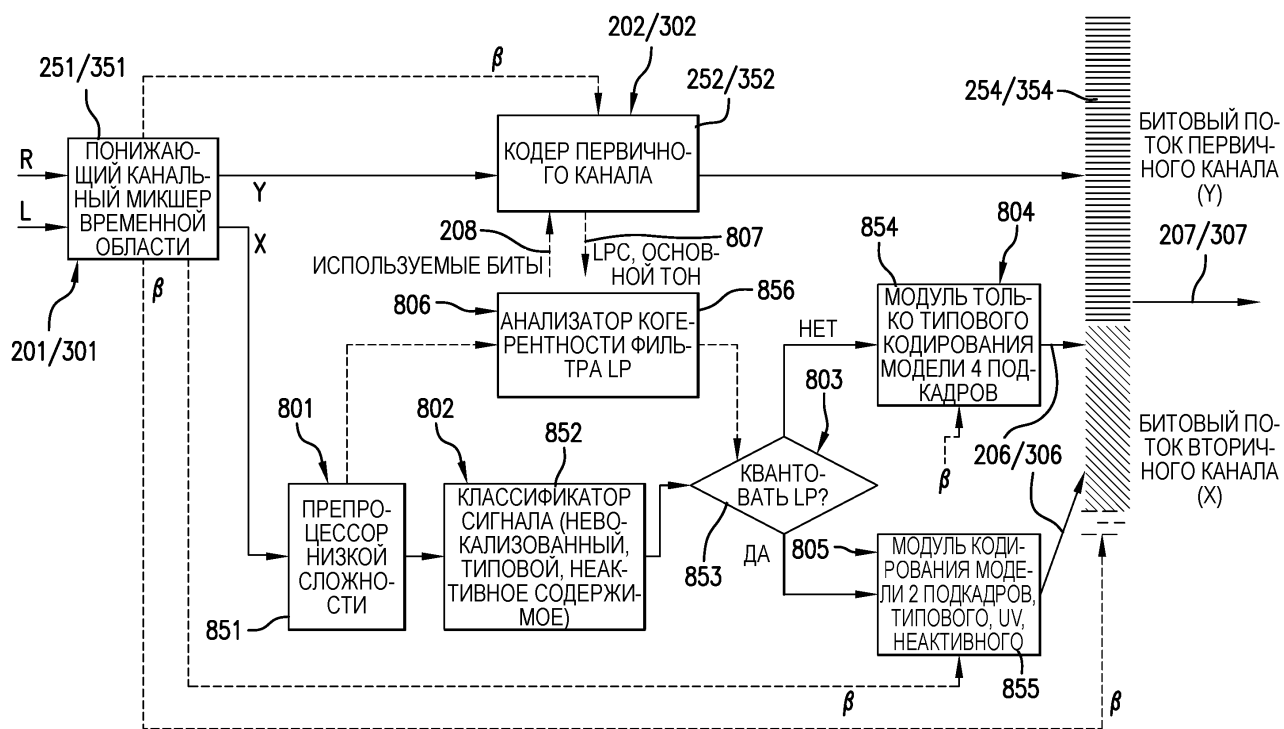
(57) Реферат:

Изобретение относится к средствам для кодирования стереофонического звука. Технический результат заключается в повышении эффективности кодирования. Микшируют с понижением левый и правый каналы стереофонического звукового сигнала для формирования первичного и вторичного каналов. Кодируют первичный канал и кодируют вторичный канал. При этом кодирование

вторичного канала содержит анализ когерентности между параметрами кодирования, вычисленными во время кодирования вторичного канала, и параметрами кодирования, вычисленными во время кодирования первичного канала, для принятия решения, являются ли параметры кодирования, вычисленные во время кодирования первичного канала, достаточно близкими к параметрам кодирования,

вычисленным во время кодирования вторичного канала, чтобы повторно использоваться во время

кодирования вторичного канала. 5 н. и 44 з.п. ф-лы, 18 ил.



ФИГ. 8

(56) (продолжение):
(20040517), vol. 3, pages 1008 - 1011.



FEDERAL SERVICE
FOR INTELLECTUAL PROPERTY

(51) Int. Cl.

G10L 19/008 (2013.01)*G10L 19/002* (2013.01)*G10L 19/06* (2013.01)*G10L 19/24* (2013.01)(12) **ABSTRACT OF INVENTION**

(52) CPC

G10L 19/002 (2020.02); G10L 19/008 (2020.02); G10L 19/06 (2020.02); G10L 19/24 (2020.02)(21)(22) Application: **2018114899, 22.09.2016**(24) Effective date for property rights:
22.09.2016Registration date:
11.08.2020

Priority:

(30) Convention priority:
25.09.2015 US 62/232,589;
14.07.2016 US 62/362,360(43) Application published: **25.10.2019 Bull. № 30**(45) Date of publication: **11.08.2020 Bull. № 23**(85) Commencement of national phase: **25.04.2018**(86) PCT application:
CA 2016/051107 (22.09.2016)(87) PCT publication:
WO 2017/049398 (30.03.2017)Mail address:
129090, Moskva, ul. B. Spasskaya, 25, str. 3, OOO
"Yuridicheskaya firma Gorodisskij i Partnery"

(72) Inventor(s):

VAILLANCOURT, Tommy (CA),
JELINEK, Milan (CA)

(73) Proprietor(s):

VOICEAGE CORPORATION (CA)(54) **METHOD AND SYSTEM FOR ENCODING A STEREO AUDIO SIGNAL USING PRIMARY CHANNEL ENCODING PARAMETERS FOR ENCODING A SECONDARY CHANNEL**

(57) Abstract:

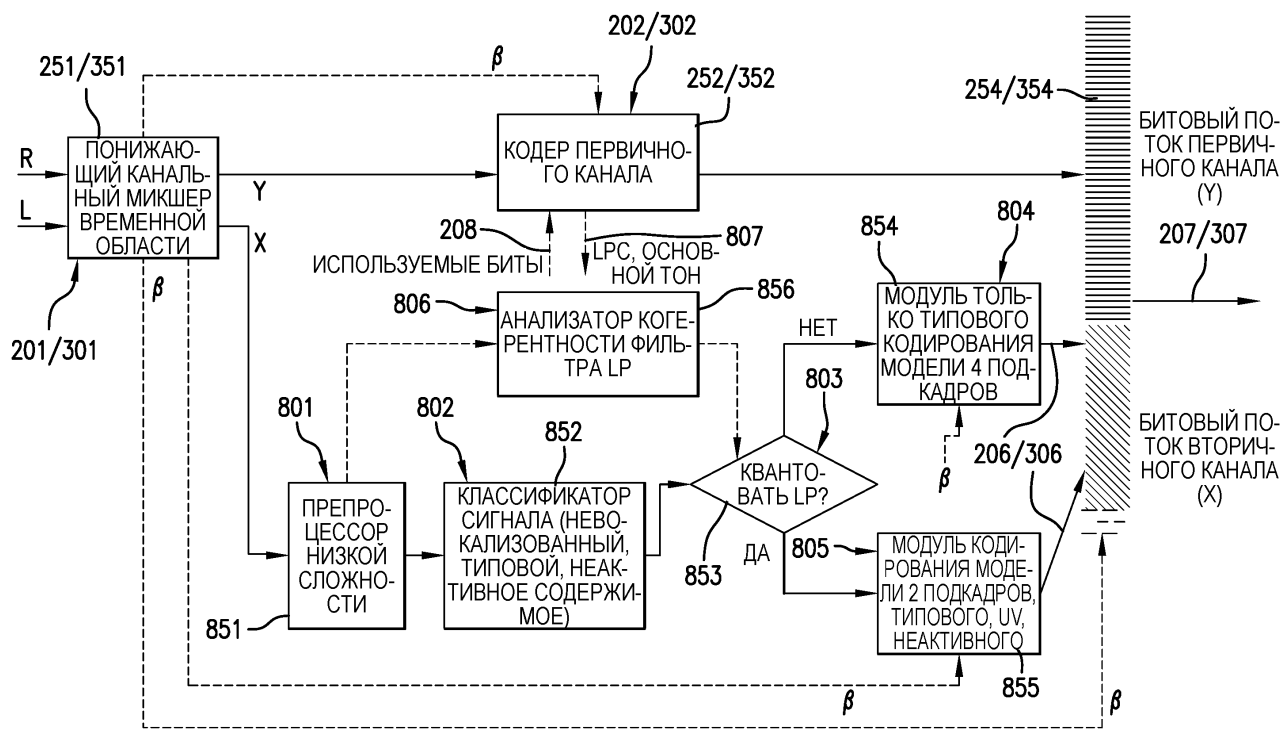
FIELD: means for encoding stereophonic sound.

SUBSTANCE: mixing with lower left and right channels of stereophonic audio signal for formation of primary and secondary channels. Primary channel is encoded and secondary channel is coded. Second channel coding comprises a coherency analysis between encoding parameters, calculated during the secondary channel encoding, and the encoding parameters

calculated during the primary channel encoding, to decide whether the encoding parameters, calculated during primary channel encoding, sufficiently close to coding parameters calculated during secondary channel coding to be reused during coding of secondary channel.

EFFECT: technical result consists in improved efficiency of coding.

49 cl, 18 dwg



ФИГ. 8

ОБЛАСТЬ ТЕХНИКИ

[0001] Настоящее раскрытие относится к кодированию стереофонического звука, в частности, но не исключительно, к кодированию стереофонической речи и/или аудио, способному создавать хорошее стереофоническое качество в сложной аудио сцене при

низкой битовой скорости и низкой задержке.

ПРЕДШЕСТВУЮЩИЙ УРОВЕНЬ ТЕХНИКИ

[0002] Исторически, разговорная телефония была реализована с трубками, имеющими только один преобразователь для вывода звука только в одно из ушей пользователя. В последнее десятилетие пользователи начали использовать свой портативный телефон вместе с наушниками, чтобы принимать звук в оба уха в основном для прослушивания музыки, а иногда и для прослушивания речи. Тем не менее, когда портативная телефонная трубка используется для передачи и приема разговорной речи, контент по-прежнему является монофоническим, но представляется в оба уха пользователя при использовании наушников.

[0003] С новейшим стандартом кодирования речи 3GPP, как описано в ссылке [1], содержание которой полностью включено в настоящий документ посредством ссылки, качество кодированного звука, например речи и/или аудио, которое передается и принимается посредством портативного телефона, было значительно улучшено. Следующим естественным шагом является передача стереофонической информации таким образом, чтобы приемник получал результат, по возможности близкий к аудио сцене реальной жизни, записанной на другом конце линии связи.

[0004] В аудиокодеках, например, как описано в ссылке [2], содержание которой полностью включено в настоящий документ посредством ссылки, обычно используется передача стереофонической информации.

[0005] Для кодеков разговорной речи, монофонический сигнал является нормой. Когда передается стереофонический сигнал, битовую скорость часто требуется удвоить, поскольку как левый, так и правый каналы кодируются с использованием монофонического кодека. Это хорошо работает в большинстве сценариев, но представляет недостатки удвоения битовой скорости и неспособности использовать любую потенциальную избыточность между двумя каналами (левым и правым каналами). Кроме того, чтобы поддерживать полную битовую скорость на приемлемом уровне, используется очень низкая битовая скорость для каждого канала, что влияет на общее качество звука.

[0006] Возможной альтернативой является использование так называемой параметрической стереофонии, как описано в ссылке [6], содержание которой полностью включено в настоящий документ посредством ссылки. Параметрическая стереосистема посылает информацию, такую как интерауральная разность времени прихода звука (ITD) или интерауральная разность интенсивности звука (IID). Последняя информация отправляется по каждому частотному диапазону, и, при низкой битовой скорости, битовый бюджет, ассоциированный со стереофонической передачей, является недостаточно высоким, чтобы позволить этим параметрам работать эффективно.

[0007] Передача коэффициента панорамирования могла бы помочь создать базовый стереоэффект при низкой битовой скорости, но такой метод не делает ничего для сохранения окружения и представляет присущие ему ограничения. Слишком быстрая адаптация коэффициента панорамирования мешает слушателю, в то время как слишком медленная адаптация коэффициента панорамирования не отражает реальное положение динамиков, что затрудняет получение хорошего качества в случае создающих помехи говорящих абонентов, или когда важна флуктуация фонового шума. В настоящее время,

кодирование разговорной стереофонической речи с подходящим качеством для всех возможных аудио сцен требует минимальной битовой скорости около 24 кбит/с для широкополосных (WB) сигналов; ниже этой битовой скорости качество речи начинает ухудшаться.

- 5 [0008] При возрастающей глобализации рабочей силы и разделении рабочих групп по всему миру необходимо улучшать связь. Например, участники телеконференции могут находиться в разных и удаленных местоположениях. Некоторые участники могут находиться в своих автомобилях, другие могут находиться в большом безэховом помещении или даже в своей гостиной. Фактически, всем участникам желательно
10 чувствовать, что они разговаривают как при живом общении. Реализация стереофонической речи, более обобщенно, стереофонического звука в портативных устройствах была бы заметным шагом в этом направлении.

КРАТКОЕ ОПИСАНИЕ СУЩНОСТИ ИЗОБРЕТЕНИЯ

- [0009] В соответствии с первым аспектом, настоящее раскрытие относится к способу
15 кодирования стереофонического звука для кодирования левого и правого каналов стереофонического звукового сигнала, содержащему понижающее микширование левого и правого каналов стереофонического звукового сигнала для формирования первичного и вторичного каналов, кодирование первичного канала и кодирование вторичного канала. Кодирование вторичного канала содержит анализ когерентности
20 между параметрами кодирования, вычисленными во время кодирования вторичного канала, и параметрами кодирования, вычисленными во время кодирования первичного канала, чтобы принимать решение, являются ли параметры кодирования, вычисленные во время кодирования первичного канала, достаточно близкими к параметрам кодирования, вычисленным во время кодирования вторичного канала, чтобы повторно
25 использоваться во время кодирования вторичного канала.

- [0010] Согласно второму аспекту, обеспечена система кодирования стереофонического звука для кодирования левого и правого каналов стереофонического звукового сигнала, содержащая понижающий микшер левого и правого каналов стереофонического звукового сигнала для формирования первичного и вторичного
30 каналов, кодер первичного канала и кодер вторичного канала. Кодер вторичного канала содержит анализатор когерентности между параметрами кодирования вторичного канала, вычисленными во время кодирования вторичного канала, и параметрами кодирования первичного канала, вычисленными во время кодирования первичного канала, чтобы принимать решение, являются ли параметры кодирования
35 первичного канала достаточно близкими к параметрам кодирования вторичного канала, чтобы повторно использоваться во время кодирования вторичного канала.

- [0011] Согласно третьему аспекту, предусмотрена система кодирования стереофонического звука для кодирования левого и правого каналов стереофонического звукового сигнала, содержащая: по меньшей мере один процессор и память, связанную
40 с процессором и содержащую не-временные инструкции, которые, при исполнении, побуждают процессор реализовывать: понижающий микшер левого и правого каналов стереофонического звукового сигнала для формирования первичного и вторичного каналов; кодер первичного канала и кодер вторичного канала; причем кодер вторичного канала содержит анализатор когерентности между параметрами кодирования
45 вторичного канала, вычисленными во время кодирования вторичного канала, и параметрами кодирования первичного канала, вычисленными во время кодирования первичного канала, чтобы принимать решение, являются ли параметры кодирования первичного канала достаточно близкими к параметрам кодирования вторичного канала,

чтобы повторно использоваться во время кодирования вторичного канала.

[0012] Еще один аспект касается системы кодирования стереофонического звука для кодирования левого и правого каналов стереофонического звукового сигнала, содержащей: по меньшей мере один процессор и память, связанную с процессором и содержащую не-временные инструкции, которые, при исполнении, побуждают процессор: выполнять понижающее микширование левого и правого каналов стереофонического звукового сигнала для формирования первичного и вторичного каналов; кодировать первичный канал с использованием кодера первичного канала и кодировать вторичный канал с использованием кодера вторичного канала и анализировать, в кодере вторичного канала, когерентность между параметрами кодирования вторичного канала, вычисленными во время кодирования вторичного канала, и параметрами кодирования первичного канала, вычисленными во время кодирования первичного канала, чтобы принимать решение, являются ли параметры кодирования первичного канала достаточно близкими к параметрам кодирования вторичного канала, чтобы повторно использоваться во время кодирования вторичного канала.

[0013] Настоящее раскрытие дополнительно относится к процессорно-читаемой памяти, содержащей не-временные инструкции, которые, при исполнении, побуждают процессор реализовывать операции описанного выше способа.

[0014] Вышеупомянутые и другие цели, преимущества и признаки способа кодирования стереофонического звука и системы для кодирования левого и правого каналов стереофонического звукового сигнала станут более очевидными после прочтения следующего неограничительного описания их иллюстративных вариантов осуществления, приведенных только в качестве примера со ссылкой на прилагаемые чертежи.

КРАТКОЕ ОПИСАНИЕ ЧЕРТЕЖЕЙ

[0015] На прилагаемых чертежах:

[0016] Фиг. 1 является блок-схемой системы обработки стереофонического звука и связи, изображающей возможный контекст реализации способа и системы кодирования стереофонического звука, как описано в нижеследующем описании;

[0017] Фиг. 2 является блок-схемой, иллюстрирующей одновременно способ и систему кодирования стереофонического звука в соответствии с первой моделью, представленной в виде схемы интегрированной стереофонии;

[0018] Фиг. 3 является блок-схемой, иллюстрирующей одновременно способ и систему кодирования стереофонического звука в соответствии с второй моделью, представленной в виде встроенной модели;

[0019] Фиг. 4 является блок-схемой, показывающей одновременно подоперации операции понижающего микширования во временной области способа кодирования стереофонического звука согласно фиг. 2 и 3 и модули канального микшера системы кодирования стереофонического звука согласно фиг. 2 и 3;

[0020] Фиг. 5 является графиком, показывающим, как линейаризованная разность долговременных корреляций отображается на коэффициент β и на коэффициент ϵ нормализации энергии;

[0021] Фиг. 6 является графиком с несколькими кривыми, показывающим разницу между использованием схемы *pca/klt* по всему кадру и использованием "косинусной" функции отображения;

[0022] Фиг. 7 является графиком с несколькими кривыми, показывающим первичный канал, вторичный канал и спектры этих первичного и вторичного каналов, являющиеся результатом применения понижающего микширования во временной области к

стереофонической выборке, которая была записана в небольшом эхо-отражающем помещении с использованием установки бинауральных микрофонов на фоне офисного шума;

[0023] Фиг. 8 является блок-схемой, иллюстрирующей одновременно способ и систему кодирования стереофонического звука, с возможной реализацией оптимизации кодирования как первичного Y, так и вторичного X каналов стереофонического звукового сигнала;

[0024] Фиг. 9 является блок-схемой, иллюстрирующей операцию анализа когерентности фильтра LP и соответствующий анализатор когерентности фильтра LP способа и системы кодирования стереофонического звука согласно фиг. 8;

[0025] Фиг. 10 является блок-схемой, иллюстрирующей одновременно способ декодирования стереофонического звука и систему декодирования стереофонического звука;

[0026] Фиг. 11 является блок-схемой, иллюстрирующей дополнительные признаки способа и системы декодирования стереофонического звука согласно фиг. 10;

[0027] Фиг. 12 является упрощенной блок-схемой примерной конфигурации аппаратных компонентов, образующих систему кодирования стереофонического звука и декодер стереофонического звука согласно настоящему раскрытию;

[0028] Фиг. 13 является блок-схемой, иллюстрирующей одновременно другие варианты осуществления подопераций операции понижающего микширования во временной области способа кодирования стереофонического звука согласно фиг. 2 и 3, и модулей канального микшера системы кодирования стереофонического звука согласно фиг. 2 и 3 с использованием коэффициента пред-адаптации для повышения стабильности стерео отображения;

[0029] Фиг. 14 является блок-схемой, иллюстрирующей одновременно операции коррекции временной задержки и модули корректора временной задержки;

[0030] Фиг. 15 является блок-схемой, иллюстрирующей одновременно альтернативный способ и систему кодирования стереофонического звука;

[0031] Фиг. 16 является блок-схемой, иллюстрирующей одновременно подоперации анализа когерентности основного тона и модули анализатора когерентности основного тона;

[0032] Фиг. 17 является блок-схемой, иллюстрирующей одновременно способ и систему стереофонического кодирования с использованием понижающего микширования во временной области с возможностью работы во временной области и в частотной области; и

[0033] Фиг. 18 является блок-схемой, иллюстрирующей одновременно другой способ и систему стереофонического кодирования с использованием понижающего микширования во временной области с возможностью работы во временной области и в частотной области.

ПОДРОБНОЕ ОПИСАНИЕ

[0034] Настоящее раскрытие относится к формированию и передаче, с низкой битовой скоростью и низкой задержкой, реалистичного представления стереофонического звукового контента, например речи и/или аудио контента, в частности, но не исключительно, из сложной аудио сцены. Сложная аудио сцена включает в себя ситуации, в которых (а) корреляция между звуковыми сигналами, которые записываются микрофонами, является низкой, (b) существует значительная флуктуация фонового шума, и/или (с) присутствует создающая помехи говорящая сторона. Примеры сложных аудио сцен содержат большой безэховый конференц-зал с конфигурацией микрофонов

А/В, небольшое эхо-отражающее помещение с бинауральными микрофонами и небольшое эхо-отражающее помещение с установкой моно/боковых микрофонов. Все эти конфигурации помещений могут включать в себя флуктуирующий фоновый шум и/или помехи от говорящих.

5 [0035] Известные кодеки стереофонического звука, такие как 3GPP AMR-WB+, как описано в ссылке [7], содержание которой полностью включено в настоящий документ посредством ссылки, являются неэффективными для кодирования звука, который не является близким к монофонической модели, особенно при низкой битовой скорости. Некоторые случаи особенно сложно кодировать с использованием существующих
10 методов стереофонии. К таким случаям относятся:

[0036] - LAAB (большое безэховое помещение с установкой А/В микрофонов);

[0037] - SEBI (небольшое эхо-отражающее помещение с установкой бинауральных микрофонов); и

15 [0038] - SEMS (небольшое эхо-отражающее помещение с установкой моно/боковых микрофонов).

[0039] Добавление флуктуирующего фонового шума и/или создающих помехи говорящих сторон приводит к тому, что эти звуковые сигналы еще труднее кодировать при низкой битовой скорости с использованием стереофонических специализированных методов, таких как параметрическая стереофония. Для кодирования таких сигналов
20 можно прибегнуть к использованию двух монофонических каналов, следовательно, удваивая битовую скорость и используемую ширину полосы сети.

[0040] Последний стандарт 3GPP EVS для разговорной речи обеспечивает диапазон битовых скоростей от 7,2 кбит/с до 96 кбит/с для широкополосной (WB) операции и от 9,6 кбит/с до 96 кбит/с для сверхширокополосной (SWB) операции. Это означает, что
25 три самые низкие удвоенные битовые скорости монофонического режима с использованием EVS составляют 14,4, 16,0 и 19,2 кбит/с для WB операции и 19,2, 26,3 и 32,8 кбит/с для SWB операции. Хотя качество речи развернутого 3GPP AMR-WB, как описано в ссылке [3], содержание которой полностью включено в настоящий документ посредством ссылки, улучшается по сравнению с его кодеком-предшественником,
30 качество кодированной речи при 7,2 кбит/с в зашумленной среде далеко от ясности, и, следовательно, можно ожидать, что качество речи двойной монофонической системы при 14,4 кбит/с также будет ограничено. При таких низких битовых скоростях, использование битовой скорости максимизировано, чтобы максимально возможное качество речи получалось как можно чаще. С использованием метода и системы
35 кодирования стереофонического звука, как описано в нижеследующем описании, минимальная полная битовая скорость передачи для контента разговорной стереофонической речи даже в случае сложных аудио сцен должна составлять около 13 кбит/с для WB и 15,0 кбит/с для SWB. При битовых скоростях, которые ниже, чем битовые скорости, используемые в двойном монофоническом подходе, качество и
40 разборчивость стереофонической речи значительно улучшаются для сложных аудио сцен.

[0041] Фиг. 1 является блок-схемой системы 100 обработки стереофонического звука и связи, изображающей возможный контекст реализации способа и системы кодирования стереофонического звука, как описано в нижеследующем описании.

45 [0042] Система 100 обработки стереофонического звука и связи согласно фиг. 1 поддерживает передачу стереофонического звукового сигнала по линии 101 связи. Линия 101 связи может содержать, например, проводную или оптоволоконную линию связи. Альтернативно, линия 101 связи может содержать, по меньшей мере

частично, радиочастотную линию связи. Радиочастотная линия связи часто поддерживает множество одновременных передач, требующих совместно используемых ресурсов ширины полосы, например, как в сотовой телефонии. Хотя не показано, линия 101 связи может быть заменена устройством памяти в реализации одиночного устройства системы 100 обработки и связи, которое записывает и сохраняет кодированный стереофонический звуковой сигнал для последующего воспроизведения.

[0043] Также со ссылкой на фиг. 1, например, пара микрофонов 102 и 122 формирует левый 103 и правый 123 каналы исходного аналогового стереофонического звукового сигнала, детектируемого, например, в сложной аудио сцене. Как указано в предшествующем описании, звуковой сигнал может содержать, в частности, но не исключительно, речь и/или аудио. Микрофоны 102 и 122 могут быть расположены в соответствии с А/В, бинауральной или моно/боковой установкой.

[0044] Левый 103 и правый 123 каналы исходного аналогового звукового сигнала подаются в аналого-цифровой (А/Д) преобразователь 104 для преобразования их в левый 105 и правый 125 каналы исходного цифрового стереофонического звукового сигнала. Левый 105 и правый 125 каналы исходного цифрового стереофонического звукового сигнала могут также быть записаны и подаваться с устройства памяти (не показано).

[0045] Кодер 106 стереофонического звука кодирует левый 105 и правый 125 каналы цифрового стереофонического звукового сигнала, тем самым создавая набор параметров кодирования, которые мультиплексируются в форме битового потока 107, доставляемого на опциональный кодер 108 с исправлением ошибок. Опциональный кодер 108 с исправлением ошибок, если присутствует, добавляет избыточность к двоичному представлению параметров кодирования в битовом потоке 107 перед передачей результирующего битового потока 111 по линии 101 связи.

[0046] На стороне приемника, опциональный декодер 109 с исправлением ошибок использует вышеупомянутую избыточную информацию в принятом цифровом битовом потоке 111 для детектирования и исправления ошибок, которые могут возникать во время передачи по линии 101 связи, создавая битовый поток 112 с принятыми параметрами кодирования. Декодер 110 стереофонического звука преобразует принятые параметры кодирования в битовый поток 112 для формирования синтезированных левого 113 и правого 133 каналов цифрового стереофонического звукового сигнала. Левый 113 и правый 133 каналы цифрового стереофонического звукового сигнала, восстановленные в декодере 110 стереофонического звука, преобразуются в синтезированные левый 114 и правый 134 каналы аналогового стереофонического звукового сигнала в цифро-аналоговом (D/A) преобразователе 115.

[0047] Синтезированные левый 114 и правый 134 каналы аналогового стереофонического звукового сигнала соответственно воспроизводятся в паре блоков 116 и 136 динамиков. В качестве альтернативы, левый 113 и правый 133 каналы цифрового стереофонического звукового сигнала от декодера 110 стереофонического звука также могут подаваться на устройство памяти (не показано) и записываться в нем.

[0048] Левый 105 и правый 125 каналы исходного цифрового стереофонического звукового сигнала согласно фиг. 1 соответствуют левому L и правому R каналам на фиг. 2, 3, 4, 8, 9, 13, 14, 15, 17 и 18. Кроме того, кодер 106 стереофонического звука на фиг. 1 соответствует системе кодирования стереофонического звука на фиг. 2, 3, 8, 15, 17 и 18.

[0049] Способ и система кодирования стереофонического звука в соответствии с

настоящим раскрытием являются двойками; предусмотрены первая и вторая модели.

[0050] Фиг. 2 является блок-схемой, иллюстрирующей одновременно способ и систему кодирования стереофонического звука в соответствии с первой моделью, представленные как интегрированная стереофоническая система, основанная на ядре EVS.

5 [0051] Со ссылкой на фиг. 2, способ кодирования стереофонического звука в соответствии с первой моделью содержит операцию 201 понижающего микширования во временной области, операцию 202 кодирования первичного канала, операцию 203 кодирования вторичного канала и операцию 204 мультиплексирования.

10 [0052] Для выполнения операции 201 понижающего микширования во временной области, канальный микшер 251 смешивает два входных стереофонических канала (правый канал R и левый канал L) для формирования первичного канала Y и вторичного канала X.

15 [0053] Для выполнения операции 203 кодирования вторичного канала, кодер 253 вторичного канала выбирает и использует минимальное количество битов (минимальную битовую скорость) для кодирования вторичного канала X с использованием одного из режимов кодирования, как определено в нижеследующем описании, и формирует соответствующий кодированный битовый поток 206 вторичного канала. Ассоциированный битовый бюджет может изменять каждый кадр в зависимости от содержимого кадра.

20 [0054] Для реализации операции 202 кодирования первичного канала используется кодер 252 первичного канала. Кодер 253 вторичного канала сигнализирует кодеру 252 первичного канала количество битов 208, используемых в текущем кадре, для кодирования вторичного канала X. Любой подходящий тип кодера может использоваться в качестве кодера 252 первичного канала. В качестве
25 неограничивающего примера, кодер 252 первичного канала может представлять собой кодер типа CELP. В этом иллюстративном варианте осуществления, кодер типа CELP первичного канала представляет собой модифицированную версию унаследованного кодера EVS, где кодер EVS модифицирован, чтобы обеспечить более высокую масштабируемость битовой скорости (битрейта), чтобы обеспечить гибкое распределение
30 битовой скорости между первичным и вторичным каналами. Таким образом, модифицированный кодер EVS сможет использовать все биты, которые не используются для кодирования вторичного канала X, для кодирования, с соответствующей битовой скоростью, первичного канала Y и формирования соответствующего кодированного битового потока 205 первичного канала.

35 [0055] Мультиплексор 254 конкатенирует битовый поток 205 первичного канала и битовый поток 206 вторичного канала для формирования мультиплексированного битового потока 207 для выполнения операции 204 мультиплексирования.

[0056] В первой модели, число битов и соответствующая битовая скорость (в битовом потоке 206), используемые для кодирования вторичного канала X, меньше, чем число
40 битов и соответствующая битовая скорость (в битовом потоке 205), используемые для кодирования первичного канала Y. Это можно рассматривать как два (2) канала с переменной битовой скоростью, причем сумма битовых скоростей двух каналов X и Y представляет собой постоянную полную битовую скорость. Этот подход может иметь разные особенности с большим или меньшим акцентом на первичный канал Y. Согласно
45 первому примеру, когда максимальный акцент делается на первичный канал Y, битовый бюджет вторичного канала X агрессивно вынуждается к минимуму. Согласно второму примеру, если меньший акцент делается на первичный канал Y, то битовый бюджет для вторичного канала X может быть сделан более постоянным, что означает, что

средняя битовая скорость вторичного канала X немного выше по сравнению с первым примером.

[0057] Напомним, что правый R и левый L каналы входного цифрового стереофонического звукового сигнала обрабатываются последовательными кадрами заданной длительности, которые могут соответствовать длительности кадров, используемых при обработке EVS. Каждый кадр содержит несколько выборок правого R и левого L каналов в зависимости от заданной длительности кадра и используемой частоты дискретизации.

[0058] Фиг. 3 является блок-схемой, иллюстрирующей одновременно способ и систему кодирования стереофонического звука в соответствии с второй моделью, представленной в виде встроенной модели.

[0059] Как показано на фиг. 3, способ кодирования стереофонического звука в соответствии с второй моделью содержит операцию 301 понижающего микширования во временной области, операцию 302 кодирования первичного канала, операцию 303 кодирования вторичного канала и операцию 304 мультиплексирования.

[0060] Для выполнения операции 301 понижающего микширования во временной области, каналный микшер 351 смешивает оба входные правый R и левый L каналы для формирования первичного канала Y и вторичного канала X.

[0061] В операции 302 кодирования первичного канала, первичный каналный кодер 352 кодирует первичный канал Y для формирования кодированного битового потока 305 первичного канала. Вновь, в качестве кодера 352 первичного канала может использоваться любой подходящий тип кодера. В качестве неограничивающего примера, кодер 352 первичного канала может представлять собой кодер типа CELP. В этом иллюстративном варианте осуществления, кодер 352 первичного канала использует, например, стандарт кодирования речи, такой как унаследованный монофонический режим кодирования EVS или режим кодирования AMR-WB-IO, что означает, что монофоническая часть битового потока 305 будет взаимодействовать с унаследованным EVS, AMR-WB-IO или унаследованным декодером AMR-WB, когда битовая скорость совместима с таким декодером. В зависимости от выбранного режима кодирования, может потребоваться некоторая регулировка первичного канала Y для обработки посредством кодера 252 первичного канала.

[0062] В операции 303 кодирования вторичного канала, кодер 353 вторичного канала кодирует вторичный канал X с меньшей битовой скоростью с использованием одного из режимов кодирования, как определено в последующем описании. Кодер 353 вторичного канала формирует кодированный битовый поток 306 вторичного канала.

[0063] Для выполнения операции 304 мультиплексирования, мультиплексор 354 конкатенирует кодированный битовый поток 305 первичного канала с кодированным битовым потоком 306 вторичного канала для формирования мультиплексированного битового потока 307. Это называется встроенной моделью, поскольку кодированный битовый поток 306 вторичного канала, ассоциированный со стерео, добавляется поверх имеющего возможность взаимодействия битового потока 305. Битовый поток 306 вторичного канала может быть удален из мультиплексированного стереофонического битового потока 307 (конкатенированных битовых потоков 305 и 306) в любой момент, что приводит к получению битового потока, декодируемого унаследованным кодеком, как описано здесь выше, в то время как пользователь новейшей версии кодека все равно сможет пользоваться полным стереофоническим декодированием.

[0064] Вышеописанные первая и вторая модели фактически близки друг к другу. Основное различие между двумя моделями заключается в возможности использовать

динамическое распределение битов между двумя каналами Y и X в первой модели, в то время как распределение битов является более ограниченным во второй модели по соображениям совместимости.

[0065] Примеры реализации и подходы, используемые для осуществления описанных выше первой и второй моделей, приведены в нижеследующем описании.

1) Понижающее микширование во временной области

[0066] Как указано в предшествующем описании, известные стереофонические модели, работающие с низкой битовой скоростью, испытывают трудности с кодированием речи, которая не близка к монофонической модели. Традиционные подходы выполняют понижающее микширование в частотной области, на каждую полосу частот, используя, например, корреляцию на каждую полосу частот, ассоциированную с анализом основных компонентов (*pca*) с использованием, например, преобразования Карунена-Лоева (Karhunen-Loève) (*klt*), для получения двух векторов, как описано в ссылках [4] и [5], содержание которых полностью включено в настоящий документ посредством ссылки. Один из этих двух векторов включает в себя все высоко коррелированное содержание, в то время как другой вектор определяет все содержание, которое не является сильно коррелированным. Наиболее известный способ кодирования речи при низких битовых скоростях использует кодек временной области, такой как кодек CELP (линейного предсказания с кодовым возбуждением), в котором известные решения частотной области непосредственно не применимы. По этой причине, хотя идея *pca/klt* на каждую полосу частот интересна, когда контент является речью, первичный канал Y должен быть преобразован обратно во временную область, и, после такого преобразования, его содержимое больше не выглядит как традиционная речь, особенно в случае описанных выше конфигураций с использованием специфической для речи модели, такой как CELP. Это приводит к снижению производительности речевого кодека. Кроме того, при низкой битовой скорости, вход речевого кодека должен быть как можно ближе к ожиданиям внутренней модели кодека.

[0067] Исходя из того, что вход речевого кодека низкой битовой скорости должен быть как можно ближе к ожидаемому речевому сигналу, был разработан первый метод. Первый метод основан на эволюции традиционной схемы *pca/klt*. В то время как традиционная схема вычисляет *pca/klt* на полосу частот, первый метод вычисляет его по всему кадру непосредственно во временной области. Это работает адекватно во время активных сегментов речи, если нет фонового шума или создающей помехи говорящей стороны. Схема *pca/klt* определяет, какой канал (левый L или правый R канал) содержит наиболее полезную информацию, этот канал отправляется в кодер первичного канала. К сожалению, схема *pca/klt* на основе кадра не надежна в присутствии фонового шума или когда два или более человека разговаривают друг с другом. Принцип схемы *pca/klt* включает в себя выбор одного входного канала (R или L) или другого, что часто приводит к резким изменениям в содержимом первичного канала, подлежащего кодированию. По меньшей мере по вышеуказанным причинам, первый метод недостаточно надежен и, соответственно, здесь представлен второй метод для преодоления недостатков первого метода и обеспечения более плавного перехода между входными каналами. Этот второй метод будет описан ниже со ссылкой на фиг. 4-9.

[0068] Со ссылкой на фиг. 4, операция понижающего микширования 201/301 временной области (фиг. 2 и 3) содержит следующие подоперации: подоперацию 401 анализа энергии, подоперацию 402 анализа тренда энергии, подоперацию 403 анализа нормализованной корреляции канала L и R, подоперацию 404 вычисления разности

долговременных (ЛТ) корреляций, подоперацию 405 преобразования разности долговременных корреляций в коэффициент β и квантования и подоперацию 406 понижающего микширования во временной области.

[0069] Имея в виду идею о том, что вход кодека звука (такого как речь и/или аудио) низкой битовой скорости должен быть как можно более однородным, подоперация 401 анализа энергии выполняется в канальном микшере 252/351 с помощью анализатора 451 энергии, чтобы сначала определить, по кадру, rms (среднеквадратичную) энергию каждого входного канала R и L, используя соотношения (1):

$$rms_L(t) = \sqrt{\frac{\sum_{i=0}^{N-1} L(i)^2}{N}}; \quad rms_R(t) = \sqrt{\frac{\sum_{i=0}^{N-1} R(i)^2}{N}} \quad (1)$$

[0070] где нижние индексы L и R обозначают соответственно левый и правый каналы, L(i) обозначает выборку i канала L, R(i) обозначает выборку i канала R, N соответствует числу выборок на кадр, и t обозначает текущий кадр.

[0071] Затем анализатор 451 энергии использует rms значения отношений (1) для определения долговременных rms значений $\overline{rms}$ для каждого канала с использованием соотношений (2):

$$\overline{rms}_L(t) = 0,6 \cdot \overline{rms}_L(t_{-1}) + 0,4 \cdot rms_L; \quad \overline{rms}_R(t) = 0,6 \cdot \overline{rms}_R(t_{-1}) + 0,4 \cdot rms_R; \quad (2)$$

[0072] где t представляет текущий кадр и t_{-1} - предыдущий кадр.

[0073] Для выполнения подоперации 402 анализа тренда энергии, анализатор 452 тренда энергии канального микшера 251/351 использует долговременные rms значения $\overline{rms}$ для определения тренда энергии в каждом канале L и R, $\overline{rms_dt}$, с использованием соотношений (3):

$$\overline{rms_dt}_L = \overline{rms}_L(t) - \overline{rms}_L(t_{-1}); \quad \overline{rms_dt}_R = \overline{rms}_R(t) - \overline{rms}_R(t_{-1}); \quad (3)$$

[0074] Тренд долговременных rms значений используется в качестве информации, которая показывает, являются ли временные события, захваченные микрофонами, постепенно затухающими, или меняют ли они каналы. Долговременные rms значения и их тренд также используются для определения скорости сходимости α разности долговременных корреляций, как будет описано ниже.

[0075] Для выполнения подоперации 403 анализа нормализованной корреляции каналов L и R, анализатор 453 нормализованной корреляции L и R вычисляет корреляцию G_{LR} для каждого из левого L и правого R каналов, нормализованных относительно монофонической версии $m(i)$ сигнала звука, такого как речь и/или аудио, в кадре t, используя отношения (4):

$$G_L(t) = \frac{\sum_{i=0}^{N-1} (L(i) \cdot m(i))}{\sum_{i=0}^{N-1} m(i)^2}, \quad G_R(t) = \frac{\sum_{i=0}^{N-1} (R(i) \cdot m(i))}{\sum_{i=0}^{N-1} m(i)^2}, \quad m(i) = \left(\frac{L(i) + R(i)}{2} \right), \quad (4)$$

[0076] где N, как уже упоминалось, соответствует числу выборок в кадре, и t обозначает текущий кадр. В данном варианте осуществления, все нормализованные корреляции и rms значения, определенные соотношениями 1-4, вычисляются во временной области для всего кадра. В другой возможной конфигурации, эти значения

могут быть вычислены в частотной области. Например, описанные здесь методы, которые адаптированы к звуковым сигналам, имеющим речевые характеристики, могут быть частью более крупной структуры, которая может переключаться между типовым способом кодирования стереофонического аудио частотной области и способом, описанным в настоящем раскрытии. В этом случае вычисление нормализованных корреляций и *rms* значений в частотной области может иметь некоторые преимущества в плане сложности или повторного использования кода.

[0077] Чтобы вычислить разность долговременных (LT) корреляций в подоперации 404, вычислитель 454 вычисляет для каждого канала L и R в текущем кадре сглаженные нормализованные корреляции с использованием соотношений (5):

$$\overline{G_L}(t) = \alpha \cdot \overline{G_L}(t_{-1}) + (1 - \alpha) \cdot G_L(t) \text{ и } \overline{G_R}(t) = \alpha \cdot \overline{G_R}(t_{-1}) + (1 - \alpha) \cdot G_R(t) \quad (5)$$

[0078] где α - упомянутая выше скорость сходимости. Наконец, вычислитель 454 определяет разность долговременных (LT) корреляций, $\overline{G_{LR}}$, используя соотношение (6):

$$\overline{G_{LR}}(t) = \overline{G_L}(t) - \overline{G_R}(t) \quad (6)$$

[0079] В одном примерном варианте осуществления, скорость сходимости α может иметь значение 0,8 или 0,5 в зависимости от долговременных энергий, вычисленных в соотношениях (2), и тренда долговременных энергий, как вычисляется в отношениях (3). Например, скорость сходимости α может иметь значение 0,8, когда долговременные энергии левого L и правого R каналов эволюционируют в одном и том же направлении, разность между разностью $\overline{G_{LR}}$ долговременных корреляций в кадре t и разностью

$\overline{G_{LR}}$ долговременных корреляций в кадре t_{-1} является низкой (ниже 0,31 для этого примерного варианта осуществления), и по меньшей мере одно из долговременных *rms* значений левого L и правого R каналов выше определенного порога (2000 в этом примерном варианте осуществления). Такие случаи означают, что оба канала L и R эволюционируют плавно, нет быстрого изменения в энергии от одного канала к другому, и по меньшей мере один канал содержит существенный уровень энергии. В противном случае, когда долговременные энергии правого R и левого L каналов эволюционируют в разных направлениях, когда разность между разностями долговременных корреляций высока, или когда оба правый R и левый L каналы имеют низкие энергии, тогда α будет установлена на 0,5, чтобы увеличить скорость адаптации разности $\overline{G_{LR}}$ долговременных корреляций.

[0080] Для выполнения подоперации 405 преобразования и квантования, после того как разность $\overline{G_{LR}}$ долговременных корреляций была надлежащим образом оценена в вычислителе 454, преобразователь и квантователь 455 преобразует эту разность в коэффициент β , который квантуется и подается в (а) кодер 252 первичного канала (фиг. 2), (b) кодер 253/353 вторичного канала (фиг. 2 и 3) и (с) мультимплексор 254/354 (фиг. 2 и 3) для передачи в декодер в мультимплексированном битовом потоке 207/307 через линию связи, такую как 101 на фиг.1.

[0081] Коэффициент β представляет два аспекта стереофонического входа, объединенные в один параметр. Во-первых, коэффициент β представляет долю или вклад каждого из правого R и левого L каналов, которые объединяются вместе для формирования первичного канала Y, а во-вторых, он также может представлять коэффициент масштабирования энергии для применения к первичному каналу Y, чтобы

получить первичный канал, который близок в энергетической области к тому, как выглядела бы монофоническая версия сигнала звука. Таким образом, в случае встроенной структуры, он позволяет самостоятельно декодировать первичный канал Y без необходимости приема вторичного битового потока 306, переносящего стереофонические параметры. Этот параметр энергии также может использоваться для повторного масштабирования энергии вторичного канала X перед его кодированием, так что глобальная энергия вторичного канала X ближе к оптимальному диапазону энергии кодера вторичного канала. Как показано на фиг. 2, информация об энергии, внутренне присутствующая в коэффициенте β , также может быть использована для улучшения распределения битов между первичным и вторичным каналами.

[0082] Квантованный коэффициент β может быть передан в декодер с использованием индекса. Так как коэффициент β может представлять как (a) соответствующие вклады левого и правого каналов в первичный канал, так и (b) коэффициент масштабирования энергии для применения к первичному каналу для получения монофонической версии сигнала звука или информации корреляции/энергии, которая помогает более эффективно распределять биты между первичным каналом Y и вторичным каналом X, то индекс, переданный в декодер, переносит два разных информационных элемента с одинаковым количеством битов.

[0083] Для получения отображения между разностью $\overline{G_{LR}}(t)$ долговременных корреляций и коэффициентом β , в этом примерном варианте осуществления, преобразователь и квантователь 455 сначала ограничивает разность $\overline{G_{LR}}(t)$ долговременных корреляций от -1,5 до 1,5, и затем линеаризует эту разность долговременных корреляций между 0 и 2, чтобы получить линеаризованную во времени разность $G'_{LR}(t)$ долговременных корреляций, как показано соотношением (7):

$$G'_{LR}(t) = \begin{cases} 0, & \overline{G_{LR}}(t) \leq -1,5 \\ \frac{2}{3} \cdot \overline{G_{LR}}(t) + 1,0, & -1,5 < \overline{G_{LR}}(t) < 1,5 \\ 2, & \overline{G_{LR}}(t) \geq 1,5 \end{cases} \quad (7)$$

[0084] В альтернативной реализации, может быть принято решение использовать только часть пространства, заполненного линеаризованной разностью $G'_{LR}(t)$ долговременных корреляций, путем дальнейшего ограничения ее значений, например, между 0,4 и 0,6. Это дополнительное ограничение привело бы к уменьшению локализации стереофонического отображения, но также к сбережению некоторого количества битов квантования. В зависимости от выбора проектирования, может учитываться этот вариант.

[0085] После линеаризации, преобразователь и квантователь 455 выполняет отображение линеаризованной разности $G'_{LR}(t)$ долговременных корреляций в "косинусную" область, используя соотношение (8):

$$\beta(t) = \frac{1}{2} \cdot \left(1 - \cos \left(\pi \cdot \frac{G'_{LR}(t)}{2} \right) \right) \quad (8)$$

[0086] Для выполнения подоперации 406 понижающего микширования во временной области, понижающий микшер 456 временной области формирует первичный канал Y и вторичный канал X в виде смеси правого R и левого L каналов, используя соотношения

(9) и (10):

$$Y(i)=R(i)\cdot(1-\beta(t))+L(i)\cdot\beta(t) \quad (9)$$

$$X(i)=L(i)\cdot(1-\beta(t))+R(i)\cdot\beta(t) \quad (10)$$

[0087] где $i=0,\dots,N-1$ - индекс выборки в кадре, и t - индекс кадра.

5 [0088] Фиг. 13 является блок-схемой, показывающей одновременно другие варианты осуществления подопераций операции 201/301 обработки понижающего микширования во временной области способа кодирования стереофонического звука согласно фиг. 2 и 3, и модулей канального микшера 251/351 системы кодирования стереофонического звука согласно фиг. 2 и 3 с использованием коэффициента предварительной адаптации
10 для повышения стабильности стереофонического отображения. В альтернативной реализации, как представлено на фиг. 13, операция 201/301 понижающего микширования во временной области содержит следующие подоперации: подоперацию 1301 анализа энергии, подоперацию 1302 анализа тренда энергии, подоперацию 1303 анализа
15 нормализованной корреляции канала L и R, подоперацию 1304 вычисления коэффициента пред-адаптации, операцию 1305 применения коэффициента пред-адаптации к нормализованным корреляциям, подоперацию 1306 вычисления разности
долговременных (LT) корреляций, подоперацию 1307 преобразования усиления в коэффициент β и квантования и подоперацию 1308 понижающего микширования во временной области.

20 [0089] Подоперации 1301, 1302 и 1303 соответственно выполняются с помощью анализатора 1351 энергии, анализатора 1352 тренда энергии и анализатора 1353 нормализованной корреляции L и R по существу таким же образом, как описано выше в отношении подопераций 401, 402 и 403 и анализаторов 451, 452 и 453 на фиг. 4.

[0090] Для выполнения подоперации 1305, канальный микшер 251/351 содержит
25 вычислитель 1355 для применения коэффициента α_r пред-адаптации непосредственно к корреляциям G_{LIR} ($G_L(t)$ и $G_R(t)$) из соотношений (4), так что их эволюция сглаживается в зависимости от энергии и характеристик обоих каналов. Если энергия сигнала мала или имеет некоторые невокализованные характеристики, то эволюция выигрыша
30 корреляции может быть более медленной.

[0091] Для выполнения подоперации 1304 вычисления коэффициента пред-адаптации, канальный микшер 251/351 содержит вычислитель 1354 коэффициента пред-адаптации, на который подаются (а) долговременные значения энергии отношений (2) левого и
35 правого каналов из анализатора 1351 энергии, (б) классификация кадров предыдущих кадров и (с) информация о речевой активности предыдущих кадров. Вычислитель 1354 коэффициента пред-адаптации вычисляет коэффициент α_r пред-адаптации, который может быть линеаризован между 0,1 и 1 в зависимости от минимальных долговременных
rms значений $\overline{rms}_{LIR}$ левого и правого каналов из анализатора 1351 с использованием соотношения (6a):

$$\alpha_r = \max(\min(M_a \cdot \min(\overline{rms}_L(t), \overline{rms}_R(t)) + B_a, 1), 0,1) \quad (11a)$$

[0092] В варианте осуществления, коэффициент M_a может иметь значение 0,0009, а коэффициент B_a - значение 0,16. В варианте, коэффициент α_r пред-адаптации может
45 быть вынужденно равным 0,15, например, если предыдущая классификация двух каналов R и L указывает на невокализованные характеристики и на активный сигнал. Флаг продолжения обнаружения голосовой активности (VAD) также может использоваться для определения того, что предыдущая часть содержания кадра была активным

сегментом.

[0093] Операция 1305 применения коэффициента пред-адаптации α_r к нормализованным корреляциям $G_{L|R}$ ($G_L(t)$ и $G_R(t)$ из соотношений (4)) левого L и правого R каналов отличается от операции 404 на фиг. 4. Вместо вычисления долговременных (LT) сглаженных нормализованных корреляций путем применения к нормализованным корреляциям $G_{L|R}$ ($G_L(t)$ и $G_R(t)$) коэффициента $(1-\alpha)$, где α является определенной выше скоростью сходимости (соотношения (5)), вычислитель 1355 применяет коэффициент α_r пред-адаптации непосредственно к нормализованным корреляциям $G_{L|R}$ ($G_L(t)$ и $G_R(t)$) левого L и правого R каналов с использованием соотношения (11b):

$$\tau_L(t) = \alpha_r \cdot G_L(t) + (1 - \alpha_r) \cdot \overline{G_L}(t) \text{ и } \tau_R(t) = \alpha_r \cdot G_R(t) + (1 - \alpha_r) \cdot \overline{G_R}(t). \quad (11b)$$

[0094] Вычислитель 1355 выводит адаптированные выигрыши (усиления) $\tau_{L|R}$ корреляций, которые подаются на вычислитель 1356 разностей долговременных (LT) корреляций. Операция понижающего микширования 201/301 во временной области (фиг. 2 и 3) содержит, в реализации согласно фиг. 13, подоперацию 1306 вычисления разности долговременных (LT) корреляций, подоперацию 1307 преобразования разности долговременных корреляций в коэффициент β и квантования и подоперацию 1358 понижающего микширования во временной области, подобно подоперациям 404, 405 и 406, соответственно, на фиг. 4.

[0095] Операция понижающего микширования 201/301 во временной области (фиг. 2 и 3) содержит, в реализации согласно фиг. 13, подоперацию 1306 вычисления разности долговременных (LT) корреляций, подоперацию 1307 преобразования разности долговременных корреляций в коэффициент β и квантования и подоперацию 1358 понижающего микширования во временной области, подобно подоперациям 404, 405 и 406, соответственно, на фиг. 4.

[0096] Подоперации 1306, 1307 и 1308 выполняются, соответственно, вычислителем 1356, преобразователем и квантователем 1357 и понижающим микшером 1358 временной области, по существу таким же образом, как описано выше в отношении подопераций 404, 405 и 406 и вычислителя 454, преобразователя и квантователя 455 и понижающего микшера 456 временной области.

[0097] Фиг. 5 показывает, как линеаризованная разность $G'_{LR}(t)$ долговременных корреляций отображается на коэффициент β и масштабирование энергии. Можно заметить, что для линеаризованной разности $G'_{LR}(t)$ долговременных корреляций, равной 1,0, означающей, что энергии/корреляции правого R и левого L каналов почти одинаковы, коэффициент β равен 0,5, и коэффициент ϵ нормализации (повторного масштабирования) энергии равен 1,0. В этой ситуации содержимое первичного канала Y в основном представляет собой монофоническую смесь, а вторичный канал X образует боковой канал. Вычисление коэффициента ϵ нормализации (повторного масштабирования) энергии описано ниже.

[0098] С другой стороны, если линеаризованная разность $G'_{LR}(t)$ долговременных корреляций равна 2, что означает, что большая часть энергии находится в левом канале L, то коэффициент β равен 1, а коэффициент нормализации (повторного масштабирования) энергии составляет 0,5, указывая, что первичный канал Y в основном содержит левый канал L в реализации интегрированной структуры или представление в уменьшенном масштабе левого канала L в реализации встроенной структуры. В этом

случае, вторичный канал X содержит правый канал R. В примерных вариантах осуществления, преобразователь и квантователь 455 или 1357 квантует коэффициент β с использованием 31 возможного элемента квантования. Квантованная версия коэффициента β представляется с использованием 5-битового индекса и, как описано выше, подается в мультиплексор для интеграции в мультиплексированный битовый поток 207/307 и передается в декодер через линию связи.

[0099] В варианте осуществления, коэффициент β также может использоваться в качестве указателя как для кодера 252/352 первичного канала, так и для кодера 253/353 вторичного канала для определения распределения битовой скорости. Например, если коэффициент β близок к 0,5, что означает, что энергии/корреляция двух (2) входных каналов с монофоническим сигналом близки друг к другу, то больше битов будет выделено вторичному каналу X, и меньше битов - первичному каналу Y, за исключением того, что если содержимое обоих каналов довольно близко, то содержимое вторичного канала будет иметь действительно низкую энергию и, вероятно, будет считаться неактивным, позволяя, таким образом, кодировать его очень малым количеством битов. С другой стороны, если коэффициент β близок к 0 или 1, то распределение битовой скорости будет поддерживать первичный канал Y.

[00100] Фиг. 6 показывает разницу между использованием вышеупомянутой схемы *pca/klt* по всему кадру (две верхние кривые на фиг. 6) по сравнению с использованием "косинусной" функции, как разложено в соотношении (8), для вычисления коэффициента β (нижняя кривая на фиг. 6). По своей природе схема *pca/klt* стремится к поиску минимума или максимума. Это хорошо работает в случае активной речи, как показано средней кривой на фиг. 6, но не очень хорошо работает для речи с фоновым шумом, поскольку она имеет тенденцию непрерывно переключаться с 0 на 1, как показано средней кривой на фиг. 6. Слишком частое переключение на экстремумы, 0 и 1, вызывает много артефактов при кодировании с низкой битовой скоростью. Потенциальное решение заключалось бы в сглаживании решений схемы *pca/klt*, но это отрицательно повлияло бы на обнаружение речевых всплесков и их корректных местоположений, тогда как "косинусная" функция согласно соотношению (8) является более эффективной в этом отношении.

[00101] На фиг. 7 показан первичный канал Y, вторичный канал X и спектры этих первичного Y и вторичного X каналов, возникающие в результате применения понижающего микширования во временной области к стереофонической выборке, которая была записана в небольшом эхо-отражающем помещении с использованием установки бинауральных микрофонов на фоне офисного шума. После операции понижающего микширования во временной области можно видеть, что оба канала по-прежнему имеют сходные формы спектра, и вторичный канал X по-прежнему имеет речеподобное временное содержимое, что позволяет использовать модель на основе речи для кодирования вторичного канала X.

[00102] Понижающее микширование во временной области, представленное в предшествующем описании, может демонстрировать некоторые проблемы в специальном случае правого R и левого L каналов, которые инвертированы по фазе. Суммирование правого R и левого L каналов для получения монофонического сигнала привело бы к тому, что правый R и левый L каналы компенсировали бы друг друга. Для решения этой возможной проблемы, в варианте осуществления, канальный микшер 251/351 сравнивает энергию монофонического сигнала с энергией как правого R, так и левого L каналов. Энергия монофонического сигнала должна быть по меньшей мере больше, чем энергия одного из правого R и левого L каналов. В противном случае, в

этом варианте осуществления, модель понижающего микширования во временной области переходит в специальный случай инвертированной фазы. В присутствии этого специального случая, коэффициент β вынужденно принимает значение 1, а вторичный канал X вынужденно кодируется с использованием типового или невокализованного режима, тем самым предотвращая неактивный режим кодирования и обеспечивая надлежащее кодирование вторичного канала X. Этот специальный случай, когда никакое изменение энергии не применяется, сигнализируется декодеру с использованием последней битовой комбинации (индексного значения), доступной для передачи коэффициента β (по существу, поскольку β квантуется с использованием 5 битов, и для квантования используется 31 элемент записи (уровень квантования), как описано выше, 32-ая возможная битовая комбинация (элемент записи или индексное значение) используется для сигнализации этого специального случая).

[00103] В альтернативной реализации, больший акцент может быть сделан на обнаружении сигналов, которые являются субоптимальными для описанных выше способов понижающего микширования и кодирования, например, в случаях несинфазных или почти несинфазных сигналов. Как только эти сигналы обнаружены, базовые методы кодирования могут быть адаптированы, если необходимо.

[00104] Обычно, для понижающего микширования во временной области, как описано здесь, когда левый L и правый R каналы входного стереофонического сигнала являются несинфазными, может произойти некоторая компенсация во время процесса понижающего микширования, что может привести к субоптимальному качеству. В приведенных выше примерах, обнаружение этих сигналов является простым, и стратегия кодирования содержит кодирование обоих каналов по отдельности. Но иногда, со специальными сигналами, такими как сигналы, которые являются несинфазными, может быть более эффективным, все еще выполнять понижающее микширование, аналогичное монофоническому/боковому варианту ($\beta=0,5$), где больший акцент делается на боковом канале. Учитывая, что некоторая специальная обработка этих сигналов может быть полезной, обнаружение таких сигналов необходимо выполнять с осторожностью. Кроме того, переход от обычной модели понижающего микширования во временной области, как описано в предшествующем описании, и модели понижающего микширования во временной области, которая имеет дело с этими специальными сигналами, может запускаться в области очень низкой энергии или в областях, где основной тон обоих каналов является нестабильным, так что переключение между двумя моделями имеет минимальный субъективный эффект.

[00105] Коррекция временной задержки (TDC) (см. корректор 1750 временной задержки на фиг. 17 и 18) между каналами L и R или метод, аналогичный тому, что описано в ссылке [8], содержание которой полностью включено в настоящий документ посредством ссылки, могут выполняться перед входом в модуль 201/301, 251/351 понижающего микширования. В таком варианте осуществления, коэффициент β может иметь смысл иной, чем было описано выше. Для такого типа реализации, при условии, что коррекция временной задержки работает, как ожидалось, коэффициент β может стать близким к 0,5, что означает, что конфигурация понижающего микширования во временной области близка к конфигурации монофонического/бокового канала. При надлежащей операции коррекции временной задержки (TDC), боковой канал может содержать сигнал, включающий в себя меньшее количество важной информации. В этом случае, битовая скорость вторичного канала X может быть минимальной, когда коэффициент β близок к 0,5. С другой стороны, если коэффициент β близок к 0 или 1, это означает, что коррекция временной задержки (TDC) не может надлежащим образом

преодолеть ситуацию рассогласования задержки, и содержимое вторичного канала X, вероятно, будет более сложным, что требует более высокой битовой скорости. Для обоих типов реализации, коэффициент β и по ассоциации коэффициент ϵ нормализации (повторного масштабирования) энергии можно использовать для улучшения

распределения битов между основным каналом Y и вторичным каналом X.

[00106] Фиг. 14 является блок-схемой, показывающей одновременно операции детектирования несинфазного сигнала и модули детектора 1450 несинфазного сигнала, образующие часть операции 201/301 понижающего микширования и канального микшера 251/351. Операции детектирования несинфазного сигнала включают в себя, как показано на фиг. 14, операцию 1401 детектирования несинфазного сигнала, операцию 1402 детектирования положения переключения и операцию 1403 выбора канального микшера для выбора между операцией 201/301 понижающего микширования во временной области и операцией 1404 специфического для несинфазного сигнала понижающего микширования во временной области. Эти операции выполняются, соответственно, с помощью детектора 1451 несинфазного сигнала, детектора 1452 положения переключения, селектора 1453 канального микшера, ранее описанного понижающего канального микшера 251/351 временной области и специфического для несинфазного сигнала понижающего канального микшера 1454 временной области.

[00107] Детектирование 1401 несинфазного сигнала основано на корреляции разомкнутого контура между первичным и вторичным каналами в предыдущих кадрах. С этой целью, детектор 1451 вычисляет в предыдущих кадрах разность $S_m(t)$ энергий между боковым сигналом $s(i)$ и монофоническим сигналом $m(i)$ с использованием соотношений (12a) и (12b):

$$S_m(t) = 10 \cdot \left(\log_{10} \left(\frac{\sqrt{\sum_{i=0}^{N-1} s(i)^2}}{N} \right) - \log_{10} \left(\frac{\sqrt{\sum_{i=0}^{N-1} m(i)^2}}{N} \right) \right) \quad (12a)$$

$$m(i) = \left(\frac{L(i) + R(i)}{2} \right) \quad \text{и} \quad s(i) = \left(\frac{L(i) - R(i)}{2} \right), \quad (12b)$$

[00108] Затем детектор 1451 вычисляет долговременную разность $\overline{S_m}(t)$ энергий между боковым и монофоническим сигналами, используя соотношение (12c):

$$\overline{S_m}(t) = \begin{cases} 0,9 \cdot \overline{S_m}(t_{-1}), & \text{для неактивного содержимого} \\ 0,9 \cdot \overline{S_m}(t_{-1}) + 0,1 \cdot S_m(t), & \text{в противном случае} \end{cases} \quad (12c)$$

[00109] где t указывает текущий кадр, t_{-1} - предыдущий кадр, и где неактивное содержимое может быть получено из флага продолжения детектора голосовой активности (VAD) или из счетчика продолжения VAD.

[00110] В дополнение к долговременной разности $\overline{S_m}(t)$ энергий между боковым и монофоническим сигналами, последняя максимальная корреляция C_{FIL} разомкнутого контура основного тона для каждого канала Y и X, как определено в разделе 5.1.10 ссылки [1], также принимается во внимание для принятия решения, когда текущая модель считается субоптимальной. $C_{P(t)}$ представляет максимальную корреляцию разомкнутого контура основного тона для первичного канала Y в предыдущем кадре,

и $C_{s(t-1)}$ - максимальную корреляцию разомкнутого контура основного тона для вторичного канала X в предыдущем кадре. Флаг F_{sub} субоптимальности вычисляется с помощью детектора 1452 положения переключения в соответствии со следующими критериями:

[00111] Если долговременная разность $\overline{S_m}(t)$ энергий между боковым и монофоническим сигналами выше некоторого порога, например, когда $\overline{S_m}(t) > 2,0$, если максимальные корреляции $C_{p(t-1)}$ и $C_{s(t-1)}$ разомкнутого контура основного тона находятся между 0,85 и 0,92, что означает, что сигналы имеют хорошую корреляцию, но они не настолько коррелированы, как был бы вокализованный сигнал, флаг субоптимальности F_{sub} устанавливается в 1, что указывает на несинфазное состояние между левым L и правым R каналами.

[00112] В противном случае, флаг субоптимальности F_{sub} устанавливается в 0, что указывает на отсутствие несинфазного состояния между левым L и правым каналами R.

[00113] Чтобы добавить некоторую стабильность в решение с флагом субоптимальности, детектор 1452 положения переключения реализует критерий относительно контура основного тона для каждого канала Y и X. Детектор 1452 положения переключения определяет, что канальный микшер 1454 будет использоваться для кодирования субоптимальных сигналов, когда, в примерном варианте осуществления, по меньшей мере три (3) последовательных экземпляра флага субоптимальности F_{sub} установлены в 1, и стабильность основного тона последнего кадра одного из первичного канала, $p_{\text{pc}(t-1)}$, или вторичного канала, $p_{\text{sc}(t-1)}$, больше, чем 64. Стабильность основного тона определяется суммой абсолютных разностей трех основных тонов разомкнутого контура, p_{0112} , как определено в 5.1.10 ссылки [1], вычисленной детектором 1452 положения переключения с использованием соотношения (12d):

$$p_{\text{pc}} = |p_1 - p_0| + |p_2 - p_1| \text{ и } p_{\text{sc}} = |p_1 - p_0| + |p_2 - p_1| \quad (12d)$$

[00114] Детектор 1452 положения переключения обеспечивает решение для селектора 1453 канального микшера, который, в свою очередь, выбирает канальный микшер 251/351 или канальный микшер 1454, соответственно. Селектор 1453 канального микшера реализует гистерезис, так что, когда выбран канальный микшер 1454, это решение выполняется до тех пор, пока не будут выполнены следующие условия: число последовательных кадров, например 20 кадров, считается оптимальным, стабильность основного тона последнего кадра одного из первичного $p_{\text{pc}(t-1)}$ или вторичного $p_{\text{sc}(t-1)}$ канала больше, чем предопределенное число, например 64, и долговременная разность $\overline{S_m}(t)$ энергий между боковым и монофоническим сигналом меньше или равна 0.

2) Динамическое кодирование между первичным и вторичным каналами

[00115] На фиг.8 показана блок-схема, иллюстрирующая одновременно способ и систему кодирования стереофонического звука, с возможной реализацией оптимизации кодирования как первичного Y, так и вторичного X каналов стереофонического звукового сигнала, такого как речь или звук.

[00116] Как показано на фиг.8, способ кодирования стереофонического звука содержит операцию 801 предварительной обработки с низкой сложностью, реализуемую препроцессором 851 низкой сложности, операцию 802 классификации сигнала,

реализуемую классификатором 852 сигнала, операцию 803 принятия решения, реализуемую модулем 853 принятия решения, операцию 804 только типового кодирования модели четырех (4) подкадров, реализуемую модулем 854 только типового кодирования модели четырех (4) подкадров, операцию 805 кодирования модели двух (2) подкадров, реализуемую модулем 855 кодирования модели двух (2) подкадров и операцию 806 анализа когерентности фильтра LP, реализуемую анализатором 856 когерентности фильтра LP.

[00117] После того как понижающее микширование 301 во временной области выполнено канальным микшером 351, в случае встроенной модели, первичный канал Y кодируется (операция 302 кодирования первичного канала) (а) с использованием в качестве кодера 352 первичного канала унаследованного кодера, такого как унаследованный кодер EVS или любой другой подходящий унаследованный звуковой кодер (следует помнить, что, как упоминалось в предшествующем описании, в качестве кодера 352 первичного канала может использоваться любой подходящий тип кодера). В случае интегрированной структуры, специализированный речевой кодек используется в качестве кодера 252 первичного канала. Специализированный речевой кодер 252 может быть кодером, основанным на переменной битовой скорости (VBR), например, модифицированной версией унаследованного кодера EVS, который был модифицирован, чтобы иметь большую масштабируемость битовой скорости, которая позволяет обрабатывать переменную битовую скорость на покадровом уровне (снова следует иметь в виду, что, как упоминалось в предшествующем описании, любой подходящий тип кодера может использоваться в качестве кодера 252 первичного канала). Это позволяет изменять минимальное количество битов, используемых для кодирования вторичного канала X, в каждом кадре и адаптировать к характеристикам кодируемого звукового сигнала. В итоге, характеристика вторичного канала X будет как можно более однородной.

[00118] Кодирование вторичного канала X, то есть более низкая энергия/корреляция с монофоническим входом, оптимизируется для использования минимальной битовой скорости, в частности, но не исключительно для речеподобного содержимого. Для этой цели кодирование вторичного канала может использовать преимущества параметров, которые уже закодированы в первичном канале Y, таких как коэффициенты фильтра LP (LPC) и/или запаздывание 807 основного тона. В частности, будет приниматься решение, как описано ниже, являются ли параметры, вычисленные во время кодирования первичного канала, достаточно близкими к соответствующим параметрам, вычисленным во время кодирования вторичного канала, чтобы повторно использоваться во время кодирования вторичного канала.

[00119] Сначала, операция 801 предварительной обработки с низкой сложностью применяется к вторичному каналу X, использующему препроцессор 851 низкой сложности, в котором в ответ на вторичный канал X вычисляются фильтр LP, детектирование голосовой активности (VAD) и основной тон разомкнутого контура. Последние вычисления могут быть реализованы, например, посредством тех, которые выполняются в унаследованном кодере EVS и описаны соответственно в разделах 5.1.9, 5.1.12 и 5.1.10 ссылки [1], содержание которой, как указано выше, полностью включено в настоящий документ посредством ссылки. Поскольку, как упоминалось в предшествующем описании, любой подходящий тип кодера может использоваться в качестве кодера 252/352 первичного канала, вышеупомянутые вычисления могут быть реализованы теми, которые выполняются в таком кодере первичного канала.

[00120] Затем характеристики сигнала вторичного канала X анализируются

классификатором 852 сигнала, чтобы классифицировать вторичный канал X как невокализованный, типовой или неактивный с использованием методов, аналогичных методам функции классификации сигнала EVS, раздел 5.1.13 той же ссылки [1]. Эти операции известны специалистам в данной области техники и для простоты могут быть
 5 взяты из стандарта 3GPP TS 26.445, v.12.0.0, но также могут использоваться альтернативные реализации.

а. Повторное использование коэффициентов фильтра LP первичного канала

[00121] Важная часть потребления битовой скорости приходится на квантование коэффициентов фильтра LP (LPC). При низкой битовой скорости, полное квантование
 10 коэффициентов фильтра LP может занимать до 25% от битового бюджета. Учитывая, что вторичный канал X часто близок по частотному содержанию к первичному каналу Y, но с наименьшим уровнем энергии, стоит проверить, можно ли повторно использовать коэффициенты фильтра LP первичного канала Y. Для этого, как показано на фиг. 8, была разработана операция 806 анализа когерентности фильтра LP, реализуемая с
 15 помощью анализатора 856 когерентности фильтра LP, в котором вычисляются и сравниваются несколько параметров для проверки возможности или невозможности повторного использования коэффициентов фильтра LP (LPC) 807 первичного канала Y.

[00122] На фиг. 9 показана блок-схема, иллюстрирующая операцию 806 анализа когерентности фильтра LP и соответствующий анализатор 856 когерентности фильтра LP для способа и системы кодирования стереофонического звука согласно фиг.8.

[00123] Операция 806 анализа когерентности фильтра LP и соответствующий анализатор 856 когерентности фильтра LP способа и системы кодирования стереофонического звука согласно фиг. 8 содержат, как показано на фиг. 9, подоперацию
 25 903 анализа фильтра LP (линейного предсказания) первичного канала, реализуемую анализатором 953 фильтра LP, подоперацию 904 взвешивания, реализуемую взвешивающим фильтром 954, подоперацию 912 анализа фильтра LP вторичного канала, реализуемую анализатором 962 фильтра LP, подоперацию 901 взвешивания, реализуемую взвешивающим фильтром 951, подоперацию 902 анализа евклидова расстояния,
 30 реализуемую анализатором 952 евклидова расстояния, подоперацию 913 фильтрации остатка, реализуемую фильтром 963 остатка, подоперацию 914 вычисления энергии остатка, реализуемую вычислителем 964 энергии остатка, подоперацию вычитания 915, реализуемую вычитателем 965, подоперацию 910 вычисления энергии звука (например, речи и/или аудио), реализуемую вычислителем 960 энергии, операцию 906 фильтрации
 35 остатка вторичного канала, реализуемую фильтром 956 остатка вторичного канала, подоперацию 907 вычисления энергии остатка, реализуемую вычислителем 957 энергии остатка, подоперацию 908 вычитания, реализуемую вычитателем 958, подоперацию 911 вычисления коэффициента усиления, реализуемую вычислителем коэффициента усиления, подоперацию 916 сравнения, реализуемую компаратором 966, подоперацию
 40 917 сравнения, реализуемую компаратором 967, подоперацию 918 принятия решения об использовании фильтра LP вторичного канала, реализуемую модулем 968 принятия решения, и подоперацию 919 принятия решения о повторном использовании фильтра LP первичного канала, реализуемую модулем 969 принятия решения.

[00124] Со ссылкой на фиг. 9, анализатор 953 фильтра LP выполняет анализ фильтра LP на первичном канале Y, в то время как анализатор 962 фильтра LP выполняет анализ
 45 фильтра LP на вторичном канале X. Анализ фильтра LP, выполняемый на каждом из первичного Y и вторичного X каналов, аналогичен анализу, описанному в разделе 5.1.9 ссылки [1].

[00125] Затем коэффициенты A_Y фильтра LP из анализатора 953 фильтра LP подаются на фильтр 956 остатка для фильтрации первого остатка r_Y вторичного канала X. Точно так же оптимальные коэффициенты A_X фильтра LP из анализатора 962 фильтра LP подаются на фильтр 963 остатка для фильтрации второго остатка r_X вторичного канала X. Фильтрация остатка с использованием коэффициентов A_Y или A_X фильтрации выполняется с использованием соотношения (11):

$$r_{Y|X}(n) = s_X(n) + \sum_{i=0}^{16} (A_{Y|X}(i) \cdot s_X(n-i)), n = 0, \dots, N-1 \quad (13)$$

[00126] где в этом примере s_X представляет вторичный канал, порядок фильтра LP равен 16, и N - число выборок в кадре (размер кадра), которое обычно равно 256 соответственно длительности кадра 20 мс при частоте дискретизации 12,8 кГц.

[00127] Вычислитель 910 вычисляет энергию E_X звукового сигнала во вторичном канале X, используя соотношение (14):

$$E_X = 10 \cdot \log_{10}(\sum_{i=0}^{N-1} s_X(i)^2), \quad (14)$$

[00128] и вычислитель 957 вычисляет энергию E_{r_Y} остатка из фильтра 956 остатка, используя соотношение (15):

$$E_{r_Y} = 10 \cdot \log_{10}(\sum_{i=0}^{N-1} r_Y(i)^2). \quad (15)$$

[00129] Вычитатель 958 вычитает энергию остатка с вычислителя 957 из звуковой энергии с вычислителя 960, чтобы получить выигрыш (усиление) G_Y предсказания.

[00130] Аналогичным образом, вычислитель 964 вычисляет энергию E_{r_X} остатка из фильтра 963 остатка, используя соотношение (16):

$$E_{r_X} = 10 \cdot \log_{10}(\sum_{i=0}^{N-1} r_X(i)^2), \quad (16)$$

[00131] и вычитатель 965 вычитает эту энергию остатка из звуковой энергии с вычислителя 960, чтобы получить усиление G_X предсказания.

[00132] Вычислитель 961 вычисляет отношение усиления G_Y/G_X . Компаратор 966 сравнивает отношение усиления G_Y/G_X с порогом τ , который равен 0,92 в примерном варианте осуществления. Если отношение G_Y/G_X меньше порога τ , то результат сравнения передается в модуль 968 принятия решения, который заставляет использовать коэффициенты фильтра LP вторичного канала для кодирования вторичного канала X.

[00133] Анализатор 952 евклидова расстояния выполняет измерение подобия фильтра LP, например, евклидова расстояния между линейными спектральными парами lsp_Y , вычисленными анализатором 953 фильтра LP в ответ на первичный канал Y, и линейными спектральными парами lsp_X , вычисленными анализатором 962 фильтра LP в ответ на вторичный канал X. Как известно специалистам в данной области техники, линейные спектральные пары lsp_Y и lsp_X представляют собой коэффициенты фильтра LP в области квантования. Анализатор 952 использует отношение (17) для определения евклидова расстояния dist:

$$dist = \sum_{i=0}^{M-1} (lsp_Y(i) - lsp_X(i))^2 \quad (17)$$

[00134] где M представляет порядок фильтра, и lsp_Y и lsp_X представляют соответственно линейные спектральные пары, вычисленные для первичного канала Y и вторичного канала X .

[00135] Перед вычислением евклидова расстояния в анализаторе 952, можно взвесить оба набора линейных спектральных пар lsp_Y и lsp_X посредством соответствующих весовых коэффициентов, так что определенные участки спектра акцентируются в большей или меньшей степени. Другие представления фильтра LP также могут использоваться для вычисления меры подобия фильтра LP.

[00136] После того как евклидово расстояние $dist$ определено, оно сравнивается с порогом σ в компараторе 967. В примерном варианте осуществления, порог σ имеет значение 0,08. Когда компаратор 966 определяет, что отношение G_Y/G_X равно или больше, чем порог τ , и компаратор 967 определяет, что евклидово расстояние $dist$ равно или больше, чем порог σ , результат сравнений передается на модуль 968 принятия решения, который вынуждает использовать коэффициенты фильтра LP вторичного канала для кодирования вторичного канала X . Когда компаратор 966 определяет, что отношение G_Y/G_X равно или больше, чем порог τ , а компаратор 967 определяет, что евклидово расстояние $dist$ меньше, чем порог σ , результат этих сравнений передается на модуль 969 принятия решения, который вынуждает повторно использовать коэффициенты фильтра LP первичного канала для кодирования вторичного канала X . В последнем случае, коэффициенты фильтра LP первичного канала повторно используются как часть кодирования вторичного канала.

[00137] Некоторые дополнительные тесты могут быть выполнены для ограничения повторного использования коэффициентов фильтра LP первичного канала для кодирования вторичного канала X в конкретных случаях, например, в случае режима невокализованного кодирования, где сигнал достаточно прост, чтобы кодировать, что еще имеется битовая скорость для кодирования также коэффициентов фильтра LP. Также возможно принудительное повторное использование коэффициентов фильтра LP первичного канала, когда уже получено очень низкое усиление остатка с коэффициентами фильтра LP вторичного канала, или когда вторичный канал X имеет очень низкий уровень энергии. Наконец, переменные τ , σ , уровень усиления остатка или очень низкий уровень энергии, при которых можно принудительно повторно использовать коэффициенты фильтра LP, могут быть адаптированы как функция доступного битового бюджета и/или как функция типа содержимого. Например, если содержимое вторичного канала считается неактивным, то даже если энергия высока, может быть принято решение повторно использовать коэффициенты фильтра LP первичного канала.

b. Кодирование при низкой битовой скорости вторичного канала

[00138] Поскольку первичный Y и вторичный X каналы могут быть комбинацией как правого R , так и левого L входных каналов, это означает, что даже если содержание энергии вторичного канала X является низким по сравнению с содержанием энергии первичного канала Y , артефакт кодирования может восприниматься после выполнения повышающего микширования каналов. Чтобы ограничить такой возможный артефакт, характеристика кодирования вторичного канала X поддерживается как можно более постоянной, чтобы ограничить любое непреднамеренное изменение энергии. Как показано на фиг. 7, содержимое вторичного канала X имеет сходные характеристики с содержимым первичного канала Y , и по этой причине была разработана модель речеподобного кодирования при низкой битовой скорости.

[00139] Со ссылкой на фиг. 8, анализатор 856 когерентности фильтра LP посылает в модуль 853 принятия решения решение повторно использовать коэффициенты фильтра LP первичного канала из модуля 969 принятия решения или решение использовать коэффициенты фильтра LP вторичного канала из модуля 968 принятия решения. Затем модуль 803 принятия решения принимает решение не квантовать коэффициенты фильтра LP вторичного канала, когда повторно используются коэффициенты фильтра LP первичного канала, и квантовать коэффициенты фильтра LP вторичного канала, когда принято решение использовать коэффициенты фильтра LP вторичного канала. В последнем случае, квантованные коэффициенты фильтра вторичного канала LP отправляются в мультиплексор 254/354 для включения в мультиплексированный битовый поток 207/307.

[00140] В операции 804 только типового кодирования модели четырех (4) подкадров и соответствующем модуле 854 только типового кодирования модели четырех (4) подкадров, чтобы поддерживать как можно более низкую битовую скорость, поиск ACELP, как описано в разделе 5.2.3.1 ссылки [1], используется только тогда, когда коэффициенты фильтра LP из первичного канала Y могут быть повторно использованы, когда вторичный канал X классифицируется как типовой посредством классификатора 852 сигнала, и когда энергия входных правого R и левого L каналов близка к центру, что означает, что энергии как правого R, так и левого L каналов близки друг к другу. Параметры кодирования, найденные во время поиска ACELP в модуле 854 только типового кодирования модели четырех (4) подкадров, затем используются для построения битового потока 206/306 вторичного канала и отправляются в мультиплексор 254/354 для включения в мультиплексированный битовый поток 207/307/

[00141] В противном случае, в операции 805 кодирования модели двух (2) подкадров и в соответствующем модуле 855 кодирования модели двух (2) подкадров используется полудиапазонная модель для кодирования вторичного канала X с типовым содержимым, когда коэффициенты фильтра LP из первичного канала Y не могут быть повторно использованы. Для неактивного и невокализованного содержимого кодируется только форма спектра.

[00142] В модуле 855 кодирования, кодирование неактивного содержимого содержит (а) кодирование усиления спектрального диапазона частотной области плюс шумовое заполнение и (b) кодирование коэффициентов фильтра LP вторичного канала, когда это необходимо, как описано соответственно в (а) разделах 5.2.3.5.7 и 5.2.3.5.11 и (b) разделе 5.2.2.1 ссылки [1]. Неактивное содержимое может быть кодировано с битовой скоростью до 1,5 кбит/с.

[00143] В модуле 855 кодирования, невокализованное кодирование вторичного канала X аналогично неактивному кодированию вторичного канала X, за исключением того, что невокализованное кодирование использует дополнительное число битов для квантования коэффициентов фильтра LP вторичного канала, которые кодированы для невокализованного вторичного канала.

[00144] Полудиапазонная модель типового кодирования построена аналогично ACELP, как описано в разделе 5.2.3.1 ссылки [1], но используется только с двумя (2) подкадрами по кадру. Таким образом, чтобы сделать это, остаток, как описано в разделе 5.2.3.1.1 ссылки [1], память адаптивной кодовой книги, как описано в разделе 5.2.3.1.4 ссылки [1], и входной вторичный канал сначала дискретизируются с понижением с коэффициентом 2. Коэффициенты фильтра LP также модифицируются для представления области с пониженной дискретизацией вместо частоты дискретизации

12,8 кГц с использованием метода, описанного в разделе 5.4.4.2 ссылки [1].

[00145] После поиска ACELP, расширение ширины полосы выполняется в частотной области возбуждения. Расширение ширины полосы сначала реплицирует энергии более низкого спектрального диапазона в более высокий диапазон. Для репликации энергий спектрального диапазона, энергия первых девяти (9) спектральных диапазонов, $G_{bd}(i)$, находится, как описано в разделе 5.2.3.5.7 ссылки [1], и последние диапазоны заполняются так, как показано в соотношении (18):

$$G_{bd}(i)=G_{bd}(16-i-1), \text{ для } i=8, \dots, 15. \quad (18)$$

[00146] Затем высокочастотное содержимое вектора возбуждения, представленного в частотной области $f_d(k)$, как описано в разделе 5.2.3.5.9 ссылки [1], заполняется с использованием частотного содержимого более низкого диапазона в соответствии с соотношением (19):

$$f_d(k)=f_d(k-P_b), \text{ для } k=128, \dots, 255, \quad (19)$$

[00147] где смещение основного тона, P_b , основано на кратном информации основного тона, как описано в разделе 5.2.3.1.4.1 ссылки [1], и преобразуется в смещение частотных бинов, как показано в соотношении (20):

$$P_b = \begin{cases} \frac{8 \cdot \left(\frac{F_s}{T}\right)}{F_r}, & \bar{T} > 64 \\ \frac{4 \cdot \left(\frac{F_s}{T}\right)}{F_r}, & \bar{T} \leq 64 \end{cases}, \quad (20)$$

[00148] где $\bar{T}$ представляет среднее значение информации декодированного основного тона на каждый подкадр, F_s представляет внутреннюю частоту дискретизации, 12,8 кГц в этом примерном варианте осуществления, и F_r - разрешение по частоте.

[00149] Параметры кодирования, найденные во время неактивного кодирования при низкой скорости, невокализованного кодирования при низкой скорости или полудиапазонного типового кодирования, выполняемого в модуле 855 кодирования модели двух (2) подкадров, затем используются для построения битового потока 206/306 вторичного канала, посылаемого в мультимплексор 254/354 для включения в мультимплексированный битовый поток 207/307.

с. Альтернативная реализация кодирования при низкой битовой скорости вторичного канала

[00150] Кодирование вторичного канала X может быть осуществлено по-другому с той же целью использования минимального количества битов при достижении наилучшего возможного качества и при сохранении постоянной характеристики. Кодирование вторичного канала X может частично управляться доступным битовым бюджетом независимо от потенциального повторного использования коэффициентов фильтра LP и информации основного тона. Кроме того, кодирование модели двух (2) подкадров (операция 805) может быть либо полудиапазонной, либо полудиапазонной. В этой альтернативной реализации кодирования при низкой битовой скорости вторичного канала, коэффициенты фильтра LP и/или информация основного тона первичного канала могут быть повторно использованы, и модель кодирования двух (2) подкадров может быть выбрана на основе битового бюджета, доступного для кодирования вторичного канала X. Кроме того, представленная ниже модель кодирования 2 подкадров была создана путем удвоения длины подкадра, вместо пониженной/повышенной дискретизации ее входных/выходных параметров.

[00151] На фиг. 15 показана блок-схема, иллюстрирующая одновременно

альтернативный способ кодирования стереофонического звука и альтернативную систему кодирования стереофонического звука. Способ и система кодирования стереофонического звука согласно фиг. 15 включают в себя несколько операций и модулей способа и системы, показанных на фиг. 8, идентифицированных с использованием одних и тех же ссылочных позиций, описание которых здесь не повторяется для краткости. Кроме того, способ кодирования стереофонического звука согласно фиг. 15 содержит операцию 1501 предварительной обработки, применяемую к первичному каналу Y до его кодирования в операции 202/302, операцию 1502 анализа когерентности основного тона, операцию 1504 принятия решения о невокализованном/неактивном сигнале, операцию 1505 принятия решения о кодировании невокализованного/неактивного сигнала и операцию 1506 принятия решения о модели 2/4 подкадров.

[00152] Подоперации 1501, 1502, 1503, 1504, 1505 и 1506 соответственно выполняются препроцессором 1551, аналогичным препроцессору 851 низкой сложности, анализатором 1552 когерентности основного тона, оценщиком 1553 распределения битов, модулем 1554 принятия решения о невокализованном/неактивном сигнале, модулем 1555 принятия решения о кодировании невокализованного/неактивного сигнала и модулем 1556 принятия решения о модели 2/4 подкадров.

[00153] Для выполнения операции 1502 анализа когерентности основного тона, на анализатор 1552 когерентности основного тона препроцессорами 851 и 1551 подаются основные тона разомкнутого контура как первичного Y, так и вторичного X каналов, соответственно $OLpitch_{pri}$ и $OLpitch_{sec}$. Анализатор 1552 когерентности основного тона согласно фиг. 15 более подробно показан на фиг. 16, которая является блок-схемой, иллюстрирующей одновременно подоперации операции 1502 анализа когерентности основного тона и модули анализатора 1552 когерентности основного тона.

[00154] Операция 1502 анализа когерентности основного тона выполняет оценку сходства основных тонов разомкнутого контура между первичным каналом Y и вторичным каналом X, чтобы принять решение, при каких условиях первичный основной тон разомкнутого контура может быть повторно использован при кодировании вторичного канала X. С этой целью, операция 1502 анализа когерентности основного тона содержит подоперацию 1601 суммирования основных тонов разомкнутого контура первичного канала, выполняемую посредством сумматора 1651 основных тонов разомкнутого контура первичного канала, и подоперацию 1602 суммирования основных тонов разомкнутого контура вторичного канала, выполняемую посредством сумматора 1652 основных тонов разомкнутого контура вторичного канала. Результат суммирования с сумматора 1652 вычитается (подоперация 1603) из результата суммирования с сумматора 1651 с использованием вычитателя 1653. Результат вычитания из подоперации 1603 обеспечивает когерентность стереофонического основного тона. В качестве неограничивающего примера, результаты суммирования в подоперациях 1601 и 1602 основаны на трех (3) предыдущих последовательных основных тонах разомкнутого контура, доступных для каждого канала Y и X. Основные тона разомкнутого контура могут быть вычислены, например, как определено в разделе 5.1.10 ссылки [1]. Когерентность S_{pc} стереофонического основного тона вычисляется в подоперациях 1601, 1602 и 1603 с использованием соотношения (21):

$$S_{pc} = \left| \sum_{i=0}^2 p_{p(i)} - \sum_{i=0}^2 p_{s(i)} \right| \quad (21)$$

[00155] где $p_{pls(i)}$ представляет основные тона разомкнутого контура первичного

канала Y и вторичного канала X , и i представляет положение основных тонов разомкнутого контуром.

[00156] Когда когерентность стереофонического сигнала ниже предопределенного порога Δ , повторное использование информации основного тона из первичного канала Y может быть разрешено, в зависимости от доступного битового бюджета, чтобы кодировать вторичный канал X . Кроме того, в зависимости от доступного битового бюджета, можно ограничить повторное использование информации основного тона для сигналов, которые имеют вокализованную характеристику как для первичного Y , так и для вторичного X каналов.

[00157] С этой целью, операция 1502 анализа когерентности основного тона содержит подоперацию 1604 принятия решения, выполняемую модулем 1654 принятия решения, который учитывает доступный битовый бюджет и характеристики звукового сигнала (указанные, например, режимами кодирования первичного и вторичного каналов). Когда модуль 1654 принятия решения обнаруживает, что доступный битовый бюджет достаточен, или звуковые сигналы как для первичного Y , так и для вторичного X каналов не имеют вокализованной характеристики, решением является кодировать информацию основного тона, относящуюся к вторичному каналу X (1605).

[00158] Когда модуль 1654 принятия решения обнаруживает, что доступный битовый бюджет является низким для цели кодирования информации основного тона вторичного канала X , или звуковые сигналы как для первичного Y , так и для вторичного X каналов имеют вокализованную характеристику, модуль принятия решения сравнивает когерентность S_{pc} основного тона стереофонического сигнала с порогом Δ . Когда битовый бюджет является низким, порог Δ устанавливается на большее значение по сравнению с тем случаем, когда битовый бюджет является более существенным (достаточным для кодирования информации основного тона вторичного канала X). Когда абсолютное значение когерентности S_{pc} основного тона стереофонического сигнала меньше или равно порогу Δ , модуль 1654 принимает решение повторно использовать информацию основного тона из первичного канала Y для кодирования вторичного канала X (1607). Когда значение когерентности S_{pc} основного тона стереофонического сигнала выше порога Δ , модуль 1654 принимает решение кодировать информацию основного тона вторичного канала X (1605).

[00159] Обеспечение того, что каналы имеют вокализованные характеристики, увеличивает вероятность плавной эволюции основного тона, тем самым уменьшая риск добавления артефактов путем повторного использования основного тона первичного канала. В качестве неограничивающего примера, когда битовый бюджет стереофонического сигнала ниже 14 кбит/с, а когерентность S_{pc} основного тона стереофонического сигнала меньше или равна 6 ($\Delta=6$), информация первичного основного тона может быть повторно использована при кодировании вторичного канала X . Согласно другому неограничивающему примеру, если битовый бюджет стереофонического сигнала выше 14 кбит/с и ниже 26 кбит/с, то как первичный канал Y , так и вторичный канал X считаются вокализованными, и когерентность S_{pc} основного тона стереофонического сигнала сравнивается с нижним порогом $\Delta=3$, что приводит к меньшей частоте повторного использования информации основного тона первичного канала Y при битовой скорости 22 кбит/с.

[00160] Со ссылкой на фиг. 15, на блок 1553 оценки распределения битов подается коэффициент β из канального микшера 251/351, решение повторно использовать коэффициенты фильтра LP первичного канала или использовать и кодировать

коэффициенты фильтра LP вторичного канала из анализатора 856 когерентности фильтра LP и информация основного тона, определенная анализатором 1552 когерентности основного тона. В зависимости от требований кодирования первичного и вторичного каналов, блок 1553 оценки распределения битов предоставляет битовый
 5 бюджет для кодирования первичного канала Y кодеру 252/352 первичного канала и битовый бюджет для кодирования вторичного канала X модулю 1556 принятия решения. В одной возможной реализации, для всего содержимого, которое не является INACTIVE (неактивным), часть полной битовой скорости распределяется вторичному каналу. Затем битовая скорость вторичного канала будет увеличена на величину, которая
 10 связана с коэффициентом ϵ нормализации энергии (повторного масштабирования), описанным ранее как:

$$B_x = B_M + (0,25 \cdot \epsilon - 0,125) \cdot (B_t - 2 \cdot B_M) \quad (21a)$$

где B_x представляет битовую скорость, распределенную вторичному каналу X, B_t
 15 представляет доступную полную стереофоническую битовую скорость, B_M представляет минимальную битовую скорость, распределенную вторичному каналу и обычно составляющую около 20% от полной стереофонической битовой скорости. Наконец, ϵ представляет вышеописанный коэффициент нормализации энергии. Следовательно, битовая скорость, распределенная первичному каналу, соответствует разности между
 20 полной стереофонической битовой скоростью и стереофонической битовой скоростью вторичного канала. В альтернативной реализации, распределение битовой скорости вторичного канала может быть описано как:

$$B_x = \begin{cases} B_M + ((15 - \epsilon_{idx}) \cdot (B_t - 2 \cdot B_M)) \cdot 0,05, & \text{если } \epsilon_{idx} < 15 \\ B_M + ((\epsilon_{idx} - 15) \cdot (B_t - 2 \cdot B_M)) \cdot 0,05, & \text{если } \epsilon_{idx} \geq 15 \end{cases} \quad (21b)$$

25 [00161] где вновь B_x представляет битовую скорость, распределенную вторичному каналу X, B_t представляет доступную полную стереофоническую битовую скорость, B_M представляет минимальную битовую скорость, распределенную вторичному каналу. Наконец, ϵ_{idx} представляет переданный индекс коэффициента нормализации энергии.
 30 Следовательно, битовая скорость, распределенная первичному каналу, соответствует разности между полной стереофонической битовой скоростью и битовой скоростью вторичного канала. Во всех случаях для содержимого INACTIVE битовая скорость вторичного канала устанавливается на минимальную необходимую битовую скорость для кодирования спектральной формы вторичного канала, дающую битовую скорость,
 35 обычно близкую к 2 кбит/с.

[00162] Между тем, классификатор 852 сигнала предоставляет классификацию сигнала вторичного канала X на модуль 1554 принятия решения. Если модуль 1554 принятия решения определяет, что звуковой сигнал является неактивным или невокализованным, модуль 1555 кодирования невокализованного/неактивного сигнала предоставляет
 40 спектральную форму вторичного канала X в мультиплексор 254/354. Альтернативно, модуль 1554 принятия решения информирует модуль 1556 принятия решения, когда звуковой сигнал не является ни неактивным, ни невокализованным. Для таких звуковых сигналов, используя битовый бюджет для кодирования вторичного канала X, модуль 1556 принятия решения определяет, имеется ли достаточное количество доступных
 45 битов для кодирования вторичного канала X с использованием модуля 854 только типового кодирования модели четырех (4) подкадров; в противном случае модуль 1556 принятия решения выбирает кодирование вторичного канала X с использованием модуля 855 кодирования модели двух (2) подкадров. Чтобы выбрать модуль только

типового кодирования модели четырех подкадров, битовый бюджет, доступный для вторичного канала, должен быть достаточно высоким для распределения по меньшей мере 40 битов для алгебраических кодовых книг, как только все остальное квантовано или использовано повторно, включая коэффициент LP и информацию основного тона и усиления.

[00163] Как будет понятно из вышеприведенного описания, в операции 804 только типового кодирования модели четырех подкадров и соответствующем модуле 854 только типового кодирования модели четырех подкадров, чтобы поддерживать битовую скорость как можно более низкой, используется поиск ACELP, как описано в разделе 5.2.3.1 ссылки [1]. В только типовом кодировании модели четырех подкадров, информация основного тона может быть повторно использована из основного канала или нет. Параметры кодирования, найденные во время поиска ACELP, затем используются в модуле 854 только типового кодирования модели четырех (4) подкадров для построения битового потока 206/306 вторичного канала и отправляются в мультимплексор 254/354 для включения в мультимплексированный битовый поток 207/307.

[00164] В альтернативной операции 805 кодирования модели двух (2) подкадров и соответствующем модуле 855 кодирования модели двух (2) подкадров, модель типового кодирования строится аналогично ACELP, как описано в разделе 5.2.3.1 ссылки [1], но она используется только с двумя (2) подкадрами на кадр. Таким образом, для этого длина подкадров увеличивается с 64 выборок до 128 выборок, сохраняя при этом внутреннюю частоту дискретизации 12,8 кГц. Если анализатор 1552 когерентности основного тона принял решение повторно использовать информацию основного тона из первичного канала Y для кодирования вторичного канала X, то вычисляется среднее значение основных тонов первых двух подкадров первичного канала Y и используется в качестве оценки основного тона для первого полукадра вторичного канала X. Аналогично, среднее значение основных тонов последних двух подкадров первичного канала Y вычисляется и используется для второго полукадра вторичного канала X. При повторном использовании из первичного канала Y, коэффициенты фильтра LP интерполируются, и интерполяция коэффициентов фильтра LP, как описано в разделе 5.2.2.1 ссылки [1], модифицируется для адаптации к схеме двух (2) подкадров путем замены первого и третьего коэффициентов интерполяции на второй и четвертый коэффициенты интерполяции.

[00165] В варианте осуществления, показанном на фиг.15, процесс принятия решения о выборе между схемами кодирования четырех (4) подкадров и двух (2) подкадров управляется битовым бюджетом, доступным для кодирования вторичного канала X. Как упоминалось ранее, битовый бюджет вторичного канала X выводится из различных элементов, таких как доступный полный битовый бюджет, коэффициент β или коэффициент ϵ нормализации энергии, наличие или отсутствие модуля коррекции временной задержки (TDC), возможность или невозможность повторного использования коэффициентов фильтра LP и/или информации основного тона из первичного канала Y.

[00166] Абсолютная минимальная битовая скорость, используемая моделью кодирования двух (2) подкадров вторичного канала X, когда коэффициенты фильтра LP и информация основного тона повторно используются из первичного канала Y, составляет около 2 кбит/с для типового сигнала, в то время как она составляет около 3,6 кбит/с для схемы кодирования четырех (4) подкадров. Для ACELP-подобного кодера, использующего модель кодирования двух (2) или четырех (4) подкадров, значительная

часть качества обусловлена количеством битов, которые могут быть распределены для поиска алгебраической кодовой книги (АСВ), как определено в разделе 5.2.3.1.5 ссылки [1].

[00167] Тогда, чтобы максимизировать качество, идея состоит в том, чтобы
 5 сравнивать битовый бюджет, доступный как для поиска алгебраической кодовой книги (АСВ) четырех (4) подкадров, так и для поиска алгебраической кодовой книги (АСВ) двух подкадров, после учета всего, что будет кодироваться. Например, если для конкретного кадра имеется 4 кбит/с (80 битов на кадр 20 мс), доступных для кодирования вторичного канала X, и коэффициент фильтра LP может быть повторно использован,
 10 когда информация основного тона должна передаваться. Тогда из 80 битов удаляется минимальное количество битов для кодирования сигнализации вторичного канала, информации основного тона вторичного канала, усиления и алгебраической кодовой книги для двух (2) подкадров и четырех (4) подкадров, чтобы получить доступный битовый бюджет для кодирования алгебраической кодовой книги. Например, модель
 15 кодирования четырех (4) подкадров выбирается, если для кодирования четырех (4) подкадров алгебраической кодовой книги доступно по меньшей мере 40 битов, в противном случае используется схема двух (2) подкадров.

3) Аппроксимация монофонического сигнала из частичного битового потока

[00168] Как описано в предшествующем описании, понижающее микширование во
 20 временной области является монофонически подходящим, что означает, что в случае встроенной структуры, где первичный канал Y кодируется унаследованным кодеком (следует иметь в виду, что, как указано в предшествующем описании, любой подходящий тип кодера может использоваться в качестве первичного канального кодера 252/352), и стереофонические биты добавляются к битовому потоку первичного канала,
 25 стереофонические биты могут быть удалены, и унаследованный декодер может создать синтез, который субъективно близок к гипотетическому монофоническому синтезу. Для этого требуется простая нормализация энергии на стороне кодера перед кодированием первичного канала Y. Путем повторного масштабирования энергии первичного канала Y до значения, достаточно близкого к энергии монофонической
 30 версии сигнала звука, декодирование первичного канала Y унаследованным декодером может стать подобным декодированию унаследованным декодером монофонической версии сигнала звука. Функция нормализации энергии непосредственно связана с линеаризованной разностью $G'_{LR}(t)$ долговременных корреляций, вычисленной с использованием соотношения (7), и вычисляется с использованием соотношения (22):
 35

$$\varepsilon = -0,485 \cdot G'_{LR}(t)^2 + 0,9765 \cdot G'_{LR}(t) + 0,5 \quad (22)$$

[00169] Уровень нормализации показан на фиг. 5. На практике, вместо использования соотношения (22), используется таблица поиска, связывающая значения ε нормализации с каждым возможным значением коэффициента β (31 значение в этом примерном
 40 варианте осуществления). Даже если этот дополнительный этап не требуется при кодировании стереофонического звукового сигнала, например речи и/или аудио, с интегрированной моделью, это может быть полезно при декодировании только монофонического сигнала без декодирования стереофонических битов.

4) Стереофоническое декодирование и повышающее микширование

[00170] На фиг. 10 показана блок-схема, иллюстрирующая одновременно способ
 45 декодирования стереофонического звука и систему декодирования стереофонического звука. На фиг. 11 показана блок-схема, иллюстрирующая дополнительные признаки способа декодирования стереофонического звука и системы декодирования

стереофонического звука согласно фиг. 10.

[00171] Способ декодирования стереофонического звука согласно фиг. 10 и 11 содержит операцию 1007 демультимплексирования, реализуемую демультимплексором 1057, операцию 1004 декодирования первичного канала, реализуемую декодером 1054 первичного канала, операцию 1005 декодирования вторичного канала, реализуемую декодером 1055 вторичного канала, и операцию 1006 повышающего микширования во временной области, реализуемую канальным повышающим микшером 1056 временной области. Операция 1005 декодирования вторичного канала содержит, как показано на фиг. 11, операцию 1101 принятия решения, реализуемую модулем 1151 принятия решения, операцию 1102 типового декодирования четырех (4) подкадров, реализуемую типовым декодером 1152 четырех (4) подкадров и операцию 1103 декодирования двух (2) подкадров типового/невокализованного/ неактивного сигнала, реализуемую декодером 1153 двух (2) кадров типового/невокализованного/неактивного сигнала.

[00172] В системе декодирования стереофонического звука, битовый поток 1001 принимается от кодера. Демультимплексор 1057 принимает битовый поток 1001 и извлекает из него параметры кодирования первичного канала Y (битовый поток 1002), параметры кодирования вторичного канала X (битовый поток 1003) и коэффициент β , подаваемые на декодер 1054 первичного канала, декодер 1055 вторичного канала и канальный повышающий микшер 1056. Как упоминалось ранее, коэффициент β используется как указатель для кодера 252/352 первичного канала и для кодера 253/353 вторичного канала для определения распределения битовой скорости, таким образом, декодер 1054 первичного канала и декодер 1055 вторичного канала оба повторно используют коэффициент β для надлежащего декодирования битового потока.

[00173] Параметры кодирования первичного канала соответствуют модели кодирования ACELP с принятой битовой скоростью и могут быть связаны с унаследованным или модифицированным кодером EVS (здесь следует иметь в виду, что, как указано в предшествующем описании, любые подходящие типы кодера могут использоваться в качестве кодера 252 первичного канала). На декодер 1054 первичного канала подается битовый поток 1002 для декодирования параметров кодирования первичного канала (codec mode₁ (режим кодека), β , LPC₁, Pitch₁ (основной тон), fixed codebook indices₁ (индексы фиксированной кодовой книги) и gains₁ (усиления), как показано на фиг. 11) с использованием способа, аналогичного раскрытому в ссылке [1], для формирования декодированного первичного канала Y' .

[00174] Параметры кодирования вторичного канала, используемые декодером 1055 вторичного канала, соответствуют модели, используемой для кодирования вторичного канала X , и могут содержать:

[00175] (a) Модель типового кодирования с повторным использованием коэффициентов фильтра LP (LPC₁) и/или других параметров кодирования (таких как, например, запаздывание основного тона Pitch₁) из первичного канала Y . На типовой декодер 1152 четырех (4) подкадров (фиг. 11) декодера 1055 вторичного канала подаются коэффициенты фильтра LP (LPC₁) и/или другие параметры кодирования (такие как, например, запаздывание основного тона Pitch₁) из первичного канала Y от декодера 1054 и/или битовый поток 1003 (β , Pitch₂, fixed codebook indices₂ и gains₂, как показано на фиг. 11), и используется способ, обратный способу в модуле 854 кодирования (фиг. 8), для получения декодированного вторичного канала X' .

[00176] (b) Другие модели кодирования могут повторно использовать или могут не

использовать коэффициенты фильтра LP (LPC_1) и/или другие параметры кодирования (такие как, например, запаздывание основного тона $Pitch_1$) из первичного канала Y, включая модель полудиапазонного типового кодирования, модель невокализованного кодирования с низкой скоростью и модель неактивного кодирования с низкой скоростью. В качестве примера, модель неактивного кодирования может повторно использовать коэффициенты LPC_1 фильтра LP первичного канала. На декодер 1153 (фиг. 11) двух (2) подкадров типового/невокализованного/ неактивного сигнала декодера 1055 вторичного канала подаются коэффициенты фильтра LP (LPC_1) и/или другие параметры кодирования (такие как, например, запаздывание основного тона $Pitch_1$) из основного канала Y и/или параметры кодирования вторичного канала из битового потока 1003 ($codec\ mode_2$, β , LPC_2 , $Pitch_2$, $fixed\ codebook\ indices_2$ и $gains_2$, как показано на фиг. 11) и используются способы, обратные способам в модуле 855 кодирования (фиг. 8) для получения декодированного вторичного канала X'.

[00177] Принятые параметры кодирования, соответствующие вторичному каналу X (битовый поток 1003), содержат информацию ($codec\ mode_2$), относящуюся к используемой модели кодирования. Модуль 1151 принятия решения использует эту информацию ($codec\ mode_2$) для определения и указания типовому декодеру 1152 четырех (4) подкадров и декодеру 1153 двух (2) подкадров типового/ невокализованного/ неактивного сигнала, какая модель кодирования должна быть использована.

[00178] В случае встроенной структуры, коэффициент β используется для извлечения индекса масштабирования энергии, который хранится в таблице поиска (не показана) на стороне декодера и используется для повторного масштабирования первичного канала Y' перед выполнением операции 1006 повышающего микширования временной области. Наконец, коэффициент β подается в канальный повышающий микшер 1056 и используется для повышающего микширования декодированных первичного Y' и вторичного X' каналов. Операция 1006 повышающего микширования во временной области выполняется как инверсия соотношений (9) и (10) понижающего микширования для получения декодированных правого R' и левого L' каналов с использованием соотношений (23) и (24):

$$L'(n) = \frac{\beta(t) \cdot Y'(n) - \beta(t) \cdot X'(n) + X'(n)}{2 \cdot \beta(t)^2 - 2 \cdot \beta(t) + 1} \quad (23)$$

$$R'(n) = \frac{-\beta(t) \cdot (Y'(n) + X'(n)) + Y'(n)}{2 \cdot \beta(t)^2 - 2 \cdot \beta(t) + 1} \quad (24)$$

[00179] где $n=0, \dots, N-1$ является индексом выборки в кадре, и t является индексом кадра.

5) Интеграция кодирования во временной области и в частотной области

[00180] Для применений настоящего метода, где используется режим кодирования в частотной области, также возможно выполнение временного понижающего микширования в частотной области, чтобы несколько снизить сложность или упростить поток данных. В таких случаях один и тот же коэффициент микширования применяется ко всем спектральным коэффициентам для сохранения преимуществ понижающего микширования во временной области. Можно заметить, что это является отклонением от применения спектральных коэффициентов на полосу частот, как в случае большинства применений понижающего микширования в частотной области. Понижающий микшер 456 может быть адаптирован для вычисления соотношений (25.1)

и (25.2):

$$F_Y(k) = F_R(k) \cdot (1 - \beta(t)) + F_L(k) \cdot \beta(t) \quad (25.1)$$

$$F_X(k) = F_L(k) \cdot (1 - \beta(t)) - F_R(k) \cdot \beta(t) \quad (25.2)$$

[00181] где $F_R(k)$ представляет частотный коэффициент k правого канала R , и, аналогично, $F_L(k)$ представляет частотный коэффициент k левого канала L . Затем первичный Y и вторичный X каналы вычисляются посредством применения обратного частотного преобразования для получения временного представления сигналов понижающего микширования.

[00182] На фиг. 17 и 18 показаны возможные реализации способа и системы стереофонического кодирования временной области с использованием понижающего микширования частотной области, способного переключаться между кодированием во временной области и частотной области первичного канала Y и вторичного канала X .

[00183] Первый вариант такого способа и системы показан на фиг. 17, которая представляет собой блок-схему, иллюстрирующую одновременно способ и систему стереофонического кодирования, использующие понижающее переключение временной области с возможностью работы во временной области и в частотной области.

[00184] На фиг. 17, способ и система стереофонического кодирования включают в себя многие ранее описанные операции и модули, описанные со ссылкой на предыдущие чертежи и обозначенные теми же ссылочными позициями. Модуль 1751 принятия решения (операция 1701 принятия решения) определяет, должны ли левый L' и правый R' каналы от корректора 1750 временной задержки кодироваться во временной области или в частотной области. Если выбрано кодирование во временной области, то способ и система стереофонического кодирования согласно фиг. 17 действуют, по существу, таким же образом, как способ и система стереофонического кодирования согласно предыдущим чертежам, например, и без ограничения, как в варианте осуществления согласно фиг. 15.

[00185] Если модуль 1751 принятия решения выбирает частотное кодирование, преобразователь 1752 времени в частоту (операция 1702 преобразования времени в частоту) преобразует левый L' и правый R' каналы в частотную область. Понижающий микшер 1753 частотной области (операция 1703 понижающего микширования в частотной области) выводит первичный Y и вторичный каналы X частотной области. Первичный канал частотной области преобразуется обратно во временную область посредством преобразователя 1754 частоты во время (операции 1704 преобразования частоты во время), и результирующий первичный канал Y временной области подается в кодер 252/352 первичного канала. Вторичный канал X частотной области от понижающего микшера 1753 частотной области обрабатывается посредством обычного параметрического кодера и/или кодера 1755 остатка (операции 1705 параметрического кодирования и/или кодирования остатка).

[00186] На фиг. 18 показана блок-схема, иллюстрирующая одновременно другой способ и систему стереофонического кодирования, использующие пониженное микширование частотной области с возможностью работы во временной области и в частотной области. На фиг. 18, способ и система стереофонического кодирования аналогичны способу и системе стереофонического кодирования согласно фиг. 17, и будут описаны только новые операции и модули.

[00187] Анализатор 1851 временной области (операция 1801 анализа временной области) заменяет ранее описанный канальный микшер 251/351 временной области

(операцию 201/301 понижающего микширования временной области). Анализатор 1851 временной области включает в себя большинство модулей согласно фиг. 4, но без понижающего микшера 456 временной области. Его роль, таким образом, в основном состоит в вычислении коэффициента β . Этот коэффициент β подается на препроцессор 851 и на преобразователи 1852 и 1853 частотной области во временную область (операции 1802 и 1803 преобразования частотной области во временную область), которые соответственно преобразуют во временную область вторичный X и первичный Y каналы частотной области, принятые из понижающего микшера 1753, для кодирования во временной области. Выходом преобразователя 1852 является, таким образом, вторичный канал X временной области, который подается в препроцессор 851, в то время как выходом преобразователя 1852 является первичный канал Y временной области, который подается как на препроцессор 1551, так и на кодер 252/352.

6) Пример конфигурации аппаратных средств

[00188] На фиг. 12 показана упрощенная блок-схема примерной конфигурации компонентов аппаратных средств, формирующих каждую из вышеописанных системы кодирования стереофонического звука и системы декодирования стереофонического звука.

[00189] Каждая из системы кодирования стереофонического звука и системы декодирования стереофонического звука может быть реализована как часть мобильного терминала в составе портативного медиаплеера или в любом подобном устройстве. Каждая из системы кодирования стереофонического звука и системы декодирования стереофонического звука (обозначенная как 1200 на фиг. 12) содержит вход 1202, выход 1204, процессор 1206 и память 1208.

[00190] Вход 1202 сконфигурирован для приема левого L и правого R каналов входного стереофонического звукового сигнала в цифровой или аналоговой форме в случае системы кодирования стереофонического звука или битового потока 1001 в случае системы декодирования стереофонического звука. Выход 1204 сконфигурирован для подачи мультимплексированного битового потока 207/307 в случае системы кодирования стереофонического звука или декодированного левого канала L' и правого канала R' в случае системы декодирования стереофонического звука. Вход 1202 и выход 1204 могут быть реализованы в общем модуле, например, в последовательном устройстве ввода/вывода.

[00191] Процессор 1206 функционально соединен с входом 1202, с выходом 1204 и с памятью 1208. Процессор 1206 реализован как один или несколько процессоров для исполнения кодовых инструкций для поддержки функций различных модулей каждой системы кодирования стереофонического звука, как показано на фиг. 2, 3, 4, 8, 9, 13, 14, 15, 16, 17 и 18 и системы декодирования стереофонического звука, как показано на фиг. 10 и 11.

[00192] Память 1208 может содержать не-временную память для хранения кодовых инструкций, исполняемых процессором 1206, в частности, процессорно-читаемую память, содержащую не-временные инструкции, которые, при исполнении, побуждают процессор реализовывать операции и модули способа и системы кодирования стереофонического звука и способа и системы декодирования стереофонического звука, как описано в настоящем раскрытии. Память 1208 может также содержать оперативную память или буфер(ы) для хранения данных промежуточной обработки от различных функций, выполняемых процессором 1206.

[00193] Специалистам в данной области техники должно быть понятно, что описание способа и системы кодирования стереофонического звука и способа и системы

декодирования стереофонического звука является только иллюстративным и не подразумевается ограничивающим каким-либо образом. Специалисты в данной области техники смогут легко предложить другие варианты осуществления, с выгодой воспользовавшись настоящим раскрытием. Кроме того, описанный способ и система кодирования стереофонического звука и способ и система декодирования стереофонического звука могут быть настроены так, чтобы предлагать полезные решения для существующих потребностей и проблем кодирования и декодирования стереофонического звука.

[00194] В интересах ясности показаны и описаны не все из обычных признаков реализаций способа и системы кодирования стереофонического звука, а также способа и системы декодирования стереофонического звука. Разумеется, будет понятно, что при разработке любой такой фактической реализации способа и системы кодирования стереофонического звука и способа и системы декодирования стереофонического звука может потребоваться множество специфических для реализации решений, чтобы достичь конкретных целей разработки, таких как соответствие ограничениям приложений, системным, сетевым и коммерческим ограничениям, и что эти конкретные цели будут варьироваться от одной реализации к другой и от одного разработчика к другому. Кроме того, следует принимать во внимание, что усилия при разработке могут быть сложными и трудоемкими, но тем не менее они будут рутинной процедурой проектирования для специалистов в области обработки звука, пользующихся преимуществом настоящего раскрытия.

[00195] В соответствии с настоящим раскрытием, модули, операции обработки и/или структуры данных, описанные в настоящем документе, могут быть реализованы с использованием различных типов операционных систем, вычислительных платформ, сетевых устройств, компьютерных программ и/или машин общего назначения. Кроме того, специалистам в данной области техники должно быть понятно, что могут использоваться также устройства менее универсального типа, такие как жестко смонтированные аппаратные устройства, программируемые пользователем вентильные матрицы (FPGA), специализированные интегральные схемы (ASIC) и т.п. Если способ, содержащий последовательность операций и подопераций, реализуется процессором, компьютером или машиной, и эти операции и подоперации могут быть сохранены в виде последовательности не-временных кодовых инструкций, процессорно-читаемых, компьютером или машиной, они могут быть сохранены на материальном (осязаемом) и/или не-временном носителе.

[00196] Модули способа и системы кодирования стереофонического звука и способа декодирования и декодера стереофонического звука, как описано в настоящем документе, могут содержать программное обеспечение, встроенное программное обеспечение, аппаратные средства или любую(ые) комбинацию(и) программного обеспечения, встроенного программного обеспечения или аппаратных средств, подходящих для целей, описанных в настоящем документе.

[00197] В способе кодирования стереофонического звука и способе декодирования стереофонического звука, как описано в настоящем документе, различные операции и подоперации могут выполняться в разных порядках, и некоторые операции и подоперации могут быть опциональными.

[00198] Хотя настоящее раскрытие было описано выше в виде неограничительных иллюстративных вариантов осуществления, эти варианты осуществления могут быть модифицированы по желанию в пределах объема приложенной формулы изобретения без отклонения от сущности и характера настоящего раскрытия.

Ссылки

Следующие ссылки упоминаются в настоящем описании, и их содержание полностью включено в настоящий документ посредством ссылки.

[1] 3GPP TS 26.445, v.12.0.0, "Codec for Enhanced Voice Services (EVS); Detailed Algorithmic Description", Sep 2014.

[2] M. Neuendorf, M. Multus, N. Rettelbach, G. Fuchs, J. Robillard, J. Lecompte, S. Wilde, S. Bayer, S. Disch, C. Helmrich, R. Lefebvre, P. Gournay, et al., "The ISO/MPEG Unified Speech and Audio Coding Standard - Consistent High Quality for All Content Types and at All Bit Rates", J. Audio Eng. Soc., vol. 61, no. 12, pp. 956-977, Dec. 2013.

[3] B. Bessette, R. Salami, R. Lefebvre, M. Jelinek, J. Rotola-Pukkila, J. Vainio, H. Mikkola, and K. Järvinen, "The Adaptive Multi-Rate Wideband Speech Codec (AMR-WB)," Special Issue of IEEE Trans. Speech and Audio Proc., Vol. 10, pp.620-636, November 2002.

[4] R.G. van der Waal & R.N.J. Veldhuis, "Subband coding of stereophonic digital audio signals", Proc. IEEE ICASSP, Vol. 5, pp. 3601-3604, April 1991.

[5] Dai Yang, Hongmei Ai, Chris Kyriakakis and C.-C. Jay Kuo, "High-Fidelity Multichannel Audio Coding With Karhunen-Loève Transform", IEEE Trans. Speech and Audio Proc., Vol. 11, No.4, pp.365-379, July 2003.

[6] J. Breebaart, S. van de Par, A. Kohlrausch and E. Schuijers, "Parametric Coding of Stereo Audio", EURASIP Journal on Applied Signal Processing, Issue 9, pp. 1305-1322, 2005.

[7] 3GPP TS 26.290 V9.0.0, "Extended Adaptive Multi-Rate - Wideband (AMR-WB+) codec; Transcoding functions (Release 9)", September 2009.

[8] Jonathan A. Gibbs, "Apparatus and method for encoding a multi-channel audio signal", US 8577045 B2.

[00199] Ниже представлено дополнительное описание, показывающее другие возможные комбинации признаков в соответствии с настоящим изобретением.

[00200] Способ кодирования стереофонического звука для кодирования левого и правого каналов стереофонического звукового сигнала содержит: понижающее микширование во временной области левого и правого каналов стереофонического звукового сигнала для формирования первичного и вторичного каналов; кодирование первичного канала и кодирование вторичного канала, причем кодирование первичного канала и кодирование вторичного канала содержит выбор первой битовой скорости для кодирования первичного канала и второй битовой скорости для кодирования вторичного канала, причем первая и вторая битовые скорости выбираются в зависимости от уровня акцентирования (предыскажения) предоставляемого первичному и вторичному каналам; кодирование вторичного канала содержит вычисление коэффициентов фильтра LP в ответ на вторичный канал и анализ когерентности между коэффициентами фильтра LP, вычисленными во время кодирования вторичного канала, и коэффициентами фильтра LP, вычисленными во время кодирования первичного канала, для принятия решения, являются ли коэффициенты фильтра LP, вычисленные во время кодирования первичного канала, достаточно близкими к коэффициентам фильтра LP, вычисленным во время кодирования вторичного канала, чтобы повторно использоваться во время кодирования вторичного канала.

[00201] Способ кодирования стереофонического звука, как описано в предыдущем абзаце, может содержать, в комбинации, по меньшей мере один из следующих признаков (a)-(l).

[00202] (a) Принятие решения о том, являются ли параметры иные, чем коэффициенты фильтра LP, и вычисленные во время кодирования первичного канала, достаточно близкими к соответствующим параметрам, вычисленным во время кодирования

вторичного канала, чтобы повторно использоваться во время кодирования вторичного канала.

[00203] (b) Кодирование вторичного канала содержит использование минимального числа битов для кодирования вторичного канала; и кодирование первичного канала содержит использование, для кодирования первичного канала, всех оставшихся битов, которые не были использованы для кодирования вторичного канала.

[00204] (c) Кодирование вторичного канала содержит использование первой фиксированной битовой скорости для кодирования первичного канала; и кодирование первичного канала содержит использование второй фиксированной битовой скорости, меньшей, чем первая битовая скорость, для кодирования вторичного канала.

[00205] (d) Сумма первой и второй битовых скоростей равна постоянной полной битовой скорости.

[00206] (e) Анализ когерентности между коэффициентами фильтра LP, вычисленными во время кодирования вторичного канала, и коэффициентами фильтра LP, вычисленными во время кодирования первичного канала, содержит: определение евклидова расстояния между первыми параметрами, представляющими коэффициенты фильтра LP, вычисленные во время кодирования первичного канала, и вторыми параметрами, представляющими коэффициенты фильтра LP, вычисленные во время кодирования вторичного канала; и сравнение евклидова расстояния с первым порогом.

[00207] (f) Анализ когерентности между коэффициентами фильтра LP, вычисленными во время кодирования вторичного канала, и коэффициентами фильтра LP, вычисленными во время кодирования первичного канала, дополнительно содержит: формирование первого остатка вторичного канала с использованием коэффициентов фильтра LP, вычисленных во время кодирования первичного канала, и формирование второго остатка вторичного канала с использованием коэффициентов фильтра LP, вычисленных во время кодирования вторичного канала; формирование первого усиления предсказания с использованием первого остатка и формирование второго усиления предсказания с использованием второго остатка; вычисление отношения между первым и вторым усилениями предсказания; сравнение отношения со вторым порогом.

[00208] (g) Анализ когерентности между коэффициентами фильтра LP, вычисленными во время кодирования вторичного канала, и коэффициентами фильтра LP, вычисленными во время кодирования первичного канала, дополнительно содержит: принятие решения, в ответ на упомянутые сравнения, являются ли коэффициенты фильтра LP, вычисленные во время кодирования первичного канала, достаточно близкими к коэффициентам фильтра LP, вычисленным во время кодирования вторичного канала, чтобы повторно использоваться во время кодирования вторичного канала.

[00209] (h) Первые и вторые параметры представляют собой линейные спектральные пары.

[00210] (i) Формирование первого усиления предсказания содержит вычисление энергии первого остатка, вычисление энергии звука во вторичном канале и вычитание энергии первого остатка из энергии звука во вторичном канале; и формирование второго усиления предсказания содержит вычисление энергии второго остатка, вычисление энергии звука во вторичном канале и вычитание энергии второго остатка из энергии звука во вторичном канале.

[00211] (j) Кодирование вторичного канала содержит классификацию вторичного канала и использование модели кодирования CELP четырех подкадров, когда вторичный канал классифицируется как типовой, и принятым решением является повторно использовать коэффициенты фильтра LP, вычисленные во время кодирования

первичного канала, чтобы кодировать вторичный канал.

[00212] (k) Кодирование вторичного канала содержит классификацию вторичного канала и использование модели кодирования с низкой скоростью двух подкадров, когда вторичный канал классифицируется как неактивный, невокализованный или
 5 типовой, и принятым решением является не использовать повторно коэффициенты фильтра LP, вычисленные во время кодирования первичного канала, чтобы кодировать вторичный канал.

[00213] (l) Энергия первичного канала масштабируется до значения, достаточно близкого к энергии монофонической версии сигнала звука, так что декодирование
 10 первичного канала унаследованным декодером аналогично декодированию унаследованным декодером монофонической версии сигнала звука.

[00214] Система кодирования стереофонического звука для кодирования левого и правого каналов стереофонического звукового сигнала содержит: понижающий микшер временной области левого и правого каналов стереофонического звукового сигнала
 15 для формирования первичного и вторичного каналов; кодер первичного канала и кодер вторичного канала, причем кодер первичного канала и кодер вторичного канала выбирают первую битовую скорость для кодирования первичного канала и вторую битовую скорость для кодирования вторичного канала, причем первая и вторая битовые скорости зависят от уровня акцентирования (предыскажения), предоставляемого
 20 первичному и вторичному каналам; кодер вторичного канала содержит анализатор фильтра LP для вычисления коэффициентов фильтра LP в ответ на вторичный канал и анализатор когерентности между коэффициентами фильтра LP вторичного канала и коэффициентами фильтра LP, вычисленными в кодере первичного канала, чтобы принять решение, являются ли коэффициенты фильтра LP первичного канала достаточно
 25 близкими к коэффициентам фильтра LP вторичного канала, чтобы повторно использоваться кодером вторичного канала.

[00215] Система кодирования стереофонического звука, как описано в предыдущем абзаце, может содержать, в комбинации, по меньшей мере один из следующих признаков (1)-(12).

30 [00216] (1) Кодер вторичного канала дополнительно принимает решение, являются ли параметры иные, чем коэффициенты фильтра LP, и вычисленные в кодере первичного канала, достаточно близкими к соответствующим параметрам, вычисленным в кодере вторичного канала, чтобы повторно использоваться кодером вторичного канала.

[00217] (2) Кодер вторичного канала использует минимальное число битов для
 35 кодирования вторичного канала, и кодер первичного канала использует, для кодирования первичного канала, все оставшиеся биты, которые не были использованы кодером вторичного канала, для кодирования вторичного канала.

[00218] (3) Кодер вторичного канала использует первую фиксированную битовую скорость для кодирования первичного канала, и кодер первичного канала использует
 40 вторую фиксированную битовую скорость, более низкую, чем первая битовая скорость, для кодирования вторичного канала.

[00219] (4) Сумма первой и второй битовых скоростей равна постоянной полной битовой скорости.

[00220] (5) Анализатор когерентности между коэффициентами фильтра LP вторичного
 45 канала и коэффициентами фильтра LP первичного канала содержит: анализатор евклидова расстояния для определения евклидова расстояния между первыми параметрами, представляющими коэффициенты фильтра LP первичного канала, и вторыми параметрами, представляющими коэффициенты LP вторичного канала; и

компаратор для сравнения евклидова расстояния с первым порогом.

[00221] (6) Анализатор когерентности между коэффициентами фильтра LP вторичного канала и коэффициентами фильтра LP первичного канала содержит: первый фильтр остатка для формирования первого остатка вторичного канала с использованием 5 коэффициентов фильтра LP первичного канала и второй фильтр остатка для формирования второго остатка вторичного канала с использованием коэффициентов LP вторичного канала; средство для формирования первого усиления предсказания с использованием первого остатка и средство для формирования второго усиления предсказания с использованием второго остатка; вычислитель отношения между первым 10 и вторым усилениями предсказания; и компаратор для сравнения отношения со вторым порогом.

[00222] (7) Анализатор когерентности между коэффициентами фильтра LP вторичного канала и коэффициентами фильтра LP первичного канала дополнительно содержит: 15 модуль принятия решения для принятия решения, в ответ на сравнения, являются ли коэффициенты фильтра LP первичного канала достаточно близкими к коэффициентам фильтра LP вторичного канала, чтобы повторно использоваться кодером вторичного канала.

[00223] (8) Первые и вторые параметры представляют собой линейные спектральные пары.

20 [00224] (9) Средство для формирования первого усиления предсказания содержит вычислитель энергии первого остатка, вычислитель энергии звука во вторичном канале и вычитатель энергии первого остатка из энергии звука во вторичном канале; и средство для формирования второго усиления предсказания содержит вычислитель энергии 25 второго остатка, вычислитель энергии звука во вторичном канале и вычитатель энергии второго остатка из энергии звука во вторичном канале.

[00225] (10) Кодер вторичного канала содержит классификатор вторичного канала и модуль кодирования, использующий модель кодирования CELP четырех подкадров, 30 когда вторичный канал классифицируется как типовой, и принятым решением является повторно использовать коэффициенты фильтра LP первичного канала, чтобы кодировать вторичный канал.

[00226] (11) Кодер вторичного канала содержит классификатор вторичного канала и модуль кодирования, использующий модель кодирования двух подкадров, когда 35 вторичный канал классифицируется как неактивный, невокализованный или типовой, и принятым решением является не использовать повторно коэффициенты фильтра LP первичного канала, чтобы кодировать вторичный канал.

[00227] (12) Предусмотрены средства для повторного масштабирования энергии 40 первичного канала до значения, достаточно близкого к энергии монофонической версии сигнала звука, так что декодирование первичного канала унаследованным декодером аналогично декодированию унаследованным декодером монофонической версии сигнала звука.

[00228] Система кодирования стереофонического звука для кодирования левого и 45 правого каналов стереофонического звукового сигнала содержит: по меньшей мере один процессор и память, связанную с процессором и содержащую не-временные инструкции, которые, при исполнении, побуждают процессор реализовывать: понижающий микшер временной области левого и правого каналов стереофонического 50 звукового сигнала для формирования первичного и вторичного каналов; кодер первичного канала и кодер вторичного канала, причем кодер первичного канала и кодер вторичного канала выбирают первую битовую скорость для кодирования

первичного канала и вторую битовую скорость для кодирования вторичного канала, причем первая и вторая битовые скорости зависят от уровня предсказания, предоставляемого первичному и вторичному каналам; кодер вторичного канала содержит анализатор фильтра LP для вычисления коэффициентов фильтра LP в ответ на вторичный канал и анализатор когерентности между коэффициентами фильтра LP вторичного канала и коэффициентами фильтра LP, вычисленными в кодере первичного канала, чтобы принять решение, являются ли коэффициенты фильтра LP первичного канала достаточно близкими к коэффициентам фильтра LP вторичного канала, чтобы повторно использоваться кодером вторичного канала.

(57) Формула изобретения

1. Способ кодирования стереофонического звука для кодирования левого и правого каналов стереофонического звукового сигнала, содержащий:

понижающее микширование левого и правого каналов стереофонического звукового сигнала для формирования первичного и вторичного каналов; и

кодирование первичного канала и кодирование вторичного канала;

при этом кодирование вторичного канала содержит анализ когерентности между параметрами кодирования, вычисленными во время кодирования вторичного канала, и параметрами кодирования, вычисленными во время кодирования первичного канала, для принятия решения, являются ли параметры кодирования, вычисленные во время кодирования первичного канала, достаточно близкими к параметрам кодирования, вычисленным во время кодирования вторичного канала, чтобы повторно использоваться во время кодирования вторичного канала.

2. Способ кодирования стереофонического звука по п. 1, в котором понижающее микширование левого и правого каналов стереофонического звукового сигнала содержит понижающее микширование во временной области левого и правого каналов стереофонического звукового сигнала, чтобы сформировать первичный и вторичный каналы.

3. Способ кодирования стереофонического звука по п. 1 или 2, в котором параметры кодирования содержат коэффициенты фильтра LP.

4. Способ кодирования стереофонического звука по любому из пп. 1-3, в котором параметры кодирования содержат информацию основного тона.

5. Способ кодирования стереофонического звука по любому из пп. 1-4, в котором кодирование первичного канала и кодирование вторичного канала содержат выбор первой битовой скорости для кодирования первичного канала и второй битовой скорости для кодирования вторичного канала, причем первая и вторая битовые скорости выбираются в зависимости от уровня предсказания, предоставляемого для первичного и вторичного каналов.

6. Способ кодирования стереофонического звука по любому из пп. 1-5, в котором: кодирование вторичного канала содержит использование минимального числа битов для кодирования вторичного канала, и кодирование первичного канала содержит использование, для кодирования первичного канала, всех оставшихся битов, которые не были использованы для кодирования вторичного канала.

7. Способ кодирования стереофонического звука по любому из пп. 1-5, в котором: кодирование первичного канала содержит использование первой фиксированной битовой скорости для кодирования первичного канала, и кодирование вторичного канала содержит использование второй фиксированной

битовой скорости, более низкой, чем первая битовая скорость, для кодирования вторичного канала.

8. Способ кодирования стереофонического звука по любому из пп. 5-7, в котором сумма первой и второй битовых скоростей равна постоянной полной битовой скорости.

5 9. Способ кодирования стереофонического звука по любому из пп. 3-8, в котором анализ когерентности между коэффициентами фильтра LP, вычисленными во время кодирования вторичного канала, и коэффициентами фильтра LP, вычисленными во время кодирования первичного канала, содержит:

10 определение евклидова расстояния между первыми параметрами, представляющими коэффициенты фильтра LP, вычисленные во время кодирования первичного канала, и вторыми параметрами, представляющими коэффициенты фильтра LP, вычисленные во время кодирования вторичного канала; и

сравнение евклидова расстояния с первым порогом.

15 10. Способ кодирования стереофонического звука по п. 9, в котором анализ когерентности между коэффициентами фильтра LP, вычисленными во время кодирования вторичного канала, и коэффициентами фильтра LP, вычисленными во время кодирования первичного канала, содержит:

20 формирование первого остатка вторичного канала с использованием коэффициентов фильтра LP, вычисленных во время кодирования первичного канала, и формирование второго остатка вторичного канала с использованием коэффициентов фильтра LP, вычисленных во время кодирования вторичного канала;

формирование первого усиления предсказания с использованием первого остатка и формирование второго усиления предсказания с использованием второго остатка;

вычисление отношения между первым и вторым усилениями предсказания;

25 сравнение упомянутого отношения со вторым порогом.

11. Способ кодирования стереофонического звука по п. 10, в котором анализ когерентности между коэффициентами фильтра LP, вычисленными во время кодирования вторичного канала, и коэффициентами фильтра LP, вычисленными во время кодирования первичного канала, содержит:

30 принятие решения, в ответ на упомянутые сравнения, являются ли коэффициенты фильтра LP, вычисленные во время кодирования первичного канала, достаточно близкими к коэффициентам фильтра LP, вычисленным во время кодирования вторичного канала, чтобы повторно использоваться во время кодирования вторичного канала.

35 12. Способ кодирования стереофонического звука по любому из пп. 9-11, в котором первые и вторые параметры представляют собой линейные спектральные пары.

13. Способ кодирования стереофонического звука по любому из пп. 10-12, в котором: формирование первого усиления предсказания содержит вычисление энергии первого остатка, вычисление энергии звука во вторичном канале и вычитание энергии первого остатка из энергии звука во вторичном канале; и

40 формирование второго усиления предсказания содержит вычисление энергии второго остатка, вычисление энергии звука во вторичном канале и вычитание энергии второго остатка из энергии звука во вторичном канале.

14. Способ кодирования стереофонического звука по любому из пп. 3-13, в котором кодирование вторичного канала содержит классификацию вторичного канала и
45 использование модели кодирования CELP четырех подкадров, когда вторичный канал классифицируется как типовой, и принятым решением является повторно использовать коэффициенты фильтра LP, вычисленные во время кодирования первичного канала, чтобы кодировать вторичный канал.

15. Способ кодирования стереофонического звука по любому из пп. 3-13, в котором кодирование вторичного канала содержит классификацию вторичного канала и использование модели кодирования с низкой скоростью двух подкадров, когда вторичный канал классифицируется как неактивный, невокализованный или типовой, и принятым решением является не использовать повторно коэффициенты фильтра LP, вычисленные во время кодирования первичного канала, чтобы кодировать вторичный канал.

16. Способ кодирования стереофонического звука по любому из пп. 4-15, в котором: анализ когерентности между информацией основного тона, вычисленной во время кодирования вторичного канала, и информацией основного тона, вычисленной во время кодирования первичного канала, содержит вычисление когерентности основных тонов разомкнутого контура первичного и вторичного каналов; и

кодирование вторичного канала содержит (а) повторное использование информации основного тона из первичного канала для кодирования вторичного канала, когда когерентность основного тона ниже или равна порогу; и (b) кодирование информации основного тона вторичного канала, когда когерентность основного тона выше, чем порог.

17. Способ кодирования стереофонического звука по п. 16, в котором вычисление когерентности основного тона разомкнутого контура первичного и вторичного каналов содержит (а) суммирование основных тонов разомкнутого контура первичного канала, (b) суммирование основных тонов разомкнутого контура вторичного канала и (c) вычитание суммы основных тонов разомкнутого контура вторичного канала из суммы основных тонов разомкнутого контура первичного канала для получения когерентности основного тона.

18. Способ кодирования стереофонического звука по п. 16 или 17, содержащий: детектирование доступного битового бюджета для кодирования информации основного тона вторичного канала;

детектирование вокализованной характеристики первичного и вторичного каналов; и

повторное использование информации основного тона первичного канала для кодирования вторичного канала, когда доступный битовый бюджет является низким для цели кодирования информации основного тона вторичного канала, когда обнаружена вокализованная характеристика первичного и вторичного каналов, и когда когерентность основного тона ниже или равна порогу.

19. Способ кодирования стереофонического звука по п. 18, содержащий установку порога на большее значение, когда доступный битовый бюджет является низким для цели кодирования информации основного тона вторичного канала и/или когда обнаружена вокализованная характеристика первичного и вторичного каналов.

20. Способ по любому из пп. 1-19, в котором, когда вторичный канал классифицирован как неактивный или невокализованный, предоставляется только спектральная форма вторичного канала для кодирования вторичного канала.

21. Способ по любому из пп. 1-20, содержащий выбор между понижающим микшированием во временной области и понижающим микшированием в частотной области.

22. Способ по любому из пп. 1-21, содержащий: преобразование левого и правого каналов из временной области в частотную область; и

понижающее микширование в частотной области левого и правого каналов частотной

области для формирования первичного и вторичного каналов частотной области.

23. Способ по п. 22, содержащий:

преобразование первичного и вторичного каналов частотной области обратно во временную область для кодирования кодером временной области.

5 24. Система кодирования стереофонического звука для кодирования левого и правого каналов стереофонического звукового сигнала, содержащая:

понижающий микшер левого и правого каналов стереофонического звукового сигнала для формирования первичного и вторичного каналов; и

кодер первичного канала и кодер вторичного канала;

10 при этом кодер вторичного канала содержит анализатор когерентности между параметрами кодирования вторичного канала, вычисленными во время кодирования вторичного канала, и параметрами кодирования первичного канала, вычисленными во время кодирования первичного канала, чтобы принять решение, являются ли
15 параметры кодирования первичного канала достаточно близкими к параметрам кодирования вторичного канала, чтобы повторно использоваться во время кодирования вторичного канала.

25. Система кодирования стереофонического звука по п. 24, в которой понижающий микшер представляет собой понижающий микшер временной области левого и правого каналов стереофонического звукового сигнала.

20 26. Система кодирования стереофонического звука по п. 24 или 25, содержащая анализатор фильтра LP для вычисления коэффициентов фильтра LP, формирующих параметры кодирования.

27. Система кодирования стереофонического звука по любому из пп. 24-26, в которой параметры кодирования содержат информацию основного тона.

25 28. Система кодирования стереофонического звука по любому из пп. 24-27, в которой кодер первичного канала и кодер вторичного канала выбирают первую битовую скорость для кодирования первичного канала и вторую битовую скорость для кодирования вторичного канала, в котором первая и вторая битовые скорости
30 выбираются в зависимости от уровня предсказания, предоставляемого первичному и вторичному каналам.

29. Система кодирования стереофонического звука по любому из пп. 24-28, в которой: кодер вторичного канала использует минимальное число битов для кодирования вторичного канала, и

35 кодер первичного канала использует, для кодирования первичного канала, все оставшиеся биты, которые не были использованы кодером вторичного канала для кодирования вторичного канала.

30. Система кодирования стереофонического звука по любому из пп. 24-29, в которой: кодер первичного канала использует первую фиксированную битовую скорость для кодирования первичного канала; и

40 кодер вторичного канала использует вторую фиксированную битовую скорость, более низкую, чем первая битовая скорость, для кодирования вторичного канала.

31. Система кодирования стереофонического звука по любому из пп. 28-30, в которой сумма первой и второй битовых скоростей равна постоянной полной битовой скорости.

32. Система кодирования стереофонического звука по любому из пп. 26-31, в которой
45 анализатор когерентности между коэффициентами фильтра LP вторичного канала и коэффициентами фильтра LP первичного канала содержит:

анализатор евклидова расстояния для определения евклидова расстояния между первыми параметрами, представляющими коэффициенты фильтра LP первичного

канала, и вторыми параметрами, представляющими коэффициенты фильтра LP вторичного канала; и

компаратор для сравнения евклидова расстояния с первым порогом.

33. Система кодирования стереофонического звука по п. 32, в которой анализатор когерентности между коэффициентами фильтра LP вторичного канала и коэффициентами фильтра LP первичного канала содержит:

первый фильтр остатка для формирования первого остатка вторичного канала с использованием коэффициентов фильтра LP первичного канала и второй фильтр остатка для формирования второго остатка вторичного канала с использованием коэффициентов фильтра LP вторичного канала;

вычислитель первого усиления предсказания, использующий первый остаток, и вычислитель второго усиления предсказания, использующий второй остаток;

вычислитель отношения между первым и вторым усилениями предсказания; и компаратор для сравнения упомянутого отношения со вторым порогом.

34. Система кодирования стереофонического звука по п. 33, в которой анализатор когерентности между коэффициентами фильтра LP вторичного канала и коэффициентами фильтра LP первичного канала дополнительно содержит:

модуль принятия решения для принятия решения, в ответ на сравнения, являются ли коэффициенты фильтра LP первичного канала достаточно близкими к коэффициентам фильтра LP вторичного канала, чтобы повторно использоваться кодером вторичного канала.

35. Система кодирования стереофонического звука по любому из пп. 32-34, в которой первые и вторые параметры представляют собой линейные спектральные пары.

36. Система кодирования стереофонического звука по любому из пп. 33-35, в которой: вычислитель первого усиления предсказания содержит вычислитель энергии первого остатка, вычислитель энергии звука во вторичном канале и вычислитель энергии первого остатка из энергии звука во вторичном канале; и

вычислитель второго усиления предсказания содержит вычислитель энергии второго остатка, вычислитель энергии звука во вторичном канале и вычислитель энергии второго остатка из энергии звука во вторичном канале.

37. Система кодирования стереофонического звука по любому из пп. 24-36, в которой кодер вторичного канала содержит классификатор вторичного канала и модуль кодирования, использующий модель кодирования CELP четырех подкадров, когда вторичный канал классифицируется как типовой и принятым решением является повторно использовать коэффициенты фильтра LP первичного канала, чтобы кодировать вторичный канал.

38. Система кодирования стереофонического звука по любому из пп. 24-36, в которой кодер вторичного канала содержит классификатор вторичного канала и модуль кодирования, использующий модель кодирования двух подкадров, когда вторичный канал классифицируется как неактивный, невокализованный или типовой, и принятым решением является не использовать повторно коэффициенты фильтра первичного канала LP, чтобы кодировать вторичный канал.

39. Система кодирования стереофонического звука по любому из пп. 27-38, в которой: анализатор когерентности основного тона вычисляет когерентность основных тонов разомкнутого контура первичного и вторичного каналов; и

кодер вторичного канала (а) повторно использует информацию основного тона из первичного канала для кодирования вторичного канала, когда когерентность основного тона ниже или равна порогу; и (б) кодирует информацию основного тона вторичного

канала, когда когерентность основного тона выше, чем порог.

40. Система кодирования стереофонического звука по п. 39, в которой, для вычисления когерентности основных тонов разомкнутого контура первичного и вторичного каналов, анализатор когерентности основного тона содержит (а) сумматор основных тонов разомкнутого контура первичного канала, (b) сумматор основных тонов разомкнутого контура вторичного канала и (с) вычитатель суммы основных тонов разомкнутого контура вторичного канала из суммы основных тонов разомкнутого контура первичного канала для получения когерентности основного тона.

41. Система кодирования стереофонического звука по п. 39 или 40, в которой: анализатор когерентности основного тона обнаруживает доступный битовый бюджет для кодирования информации основного тона вторичного канала и обнаруживает вокализованную характеристику первичного и вторичного каналов; и

кодер вторичного канала повторно использует информацию основного тона первичного канала для кодирования вторичного канала, когда доступный битовый бюджет является низким для цели кодирования информации основного тона вторичного канала, когда обнаружена вокализованная характеристика первичного и вторичного каналов, и когда когерентность основного тона ниже или равна порогу.

42. Система кодирования стереофонического звука по п. 41, содержащая средство для установки порога на большее значение, когда доступный битовый бюджет является низким для цели кодирования информации основного тона вторичного канала, и/или когда обнаружена вокализованная характеристика первичного и вторичного канала.

43. Система по любому из пп. 24-42, в которой, когда вторичный канал классифицируется как неактивный или невокализованный, кодер вторичного канала предоставляет только спектральную форму вторичного канала для кодирования вторичного канала.

44. Система по любому из пп. 24-42, в которой понижающий канальный микшер выбирает между понижающим микшированием во временной области и понижающим микшированием в частотной области.

45. Система по любому из пп. 24-42 и 44, содержащая: преобразователь левого и правого каналов из временной области в частотную область;

при этом канальный понижающий микшер микширует левый и правый каналы частотной области для формирования первичного и вторичного каналов частотной области.

46. Система по п. 45, содержащая:

преобразователь первичного и вторичного каналов частотной области обратно во временную область для кодирования кодером временной области.

47. Система кодирования стереофонического звука для кодирования левого и правого каналов стереофонического звукового сигнала, содержащая:

по меньшей мере один процессор; и

память, связанную с процессором и содержащую не-временные инструкции, которые, при исполнении, побуждают процессор реализовывать:

понижающий микшер левого и правого каналов стереофонического звукового сигнала для формирования первичного и вторичного каналов; и

кодер первичного канала и кодер вторичного канала;

при этом кодер вторичного канала содержит анализатор когерентности между параметрами кодирования вторичного канала, вычисленными во время кодирования вторичного канала, и параметрами кодирования первичного канала, вычисленными

во время кодирования первичного канала, чтобы принимать решение, являются ли параметры кодирования первичного канала достаточно близкими к параметрам кодирования вторичного канала, чтобы повторно использоваться во время кодирования вторичного канала.

5 48. Система кодирования стереофонического звука для кодирования левого и правого каналов стереофонического звукового сигнала, содержащая:

по меньшей мере один процессор; и

память, связанную с процессором и содержащую не-временные инструкции, которые, при исполнении, побуждают процессор:

10 выполнять понижающее микширование левого и правого каналов стереофонического звука для формирования первичного и вторичного каналов;

кодировать первичный канал с использованием кодера первичного канала и кодировать вторичный канал с использованием кодера вторичного канала; и

15 анализировать, в кодере вторичного канала, когерентность между параметрами кодирования вторичного канала, вычисленными во время кодирования вторичного канала, и параметрами кодирования первичного канала, вычисленными во время кодирования первичного канала, для принятия решения, являются ли параметры кодирования первичного канала достаточно близкими к параметрам кодирования вторичного канала, чтобы повторно использоваться во время кодирования вторичного
20 канала.

49. Процессорно-читаемая память, содержащая не-временные инструкции, которые, при исполнении, побуждают процессор реализовывать операции способа по любому одному из пп. 1-23.

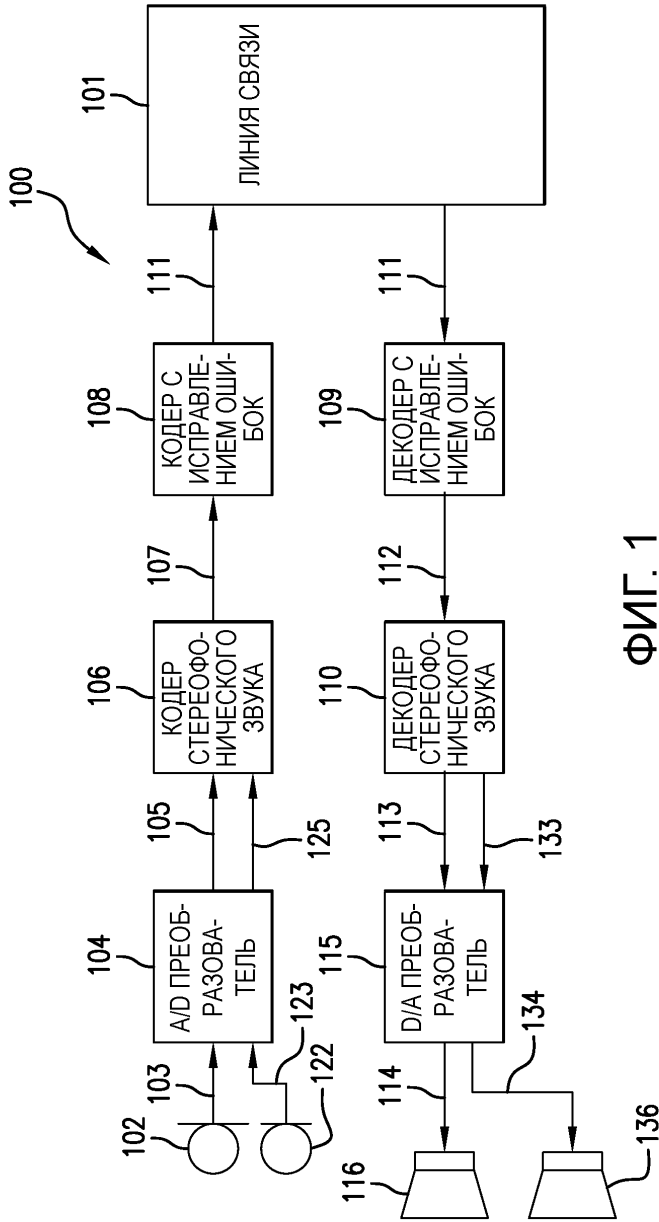
25

30

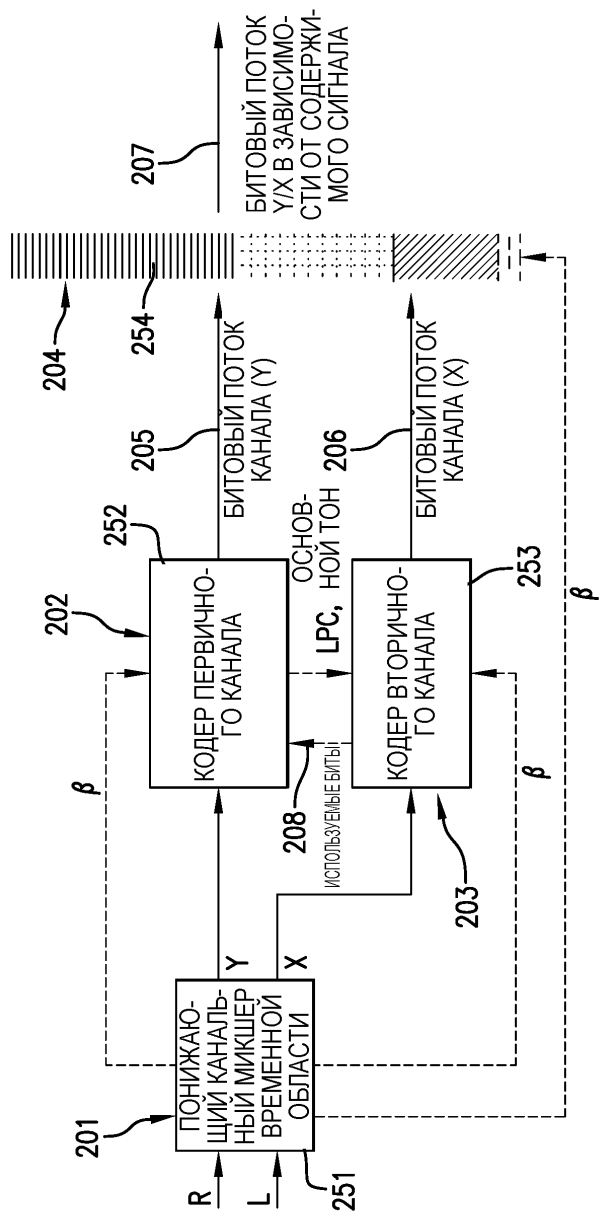
35

40

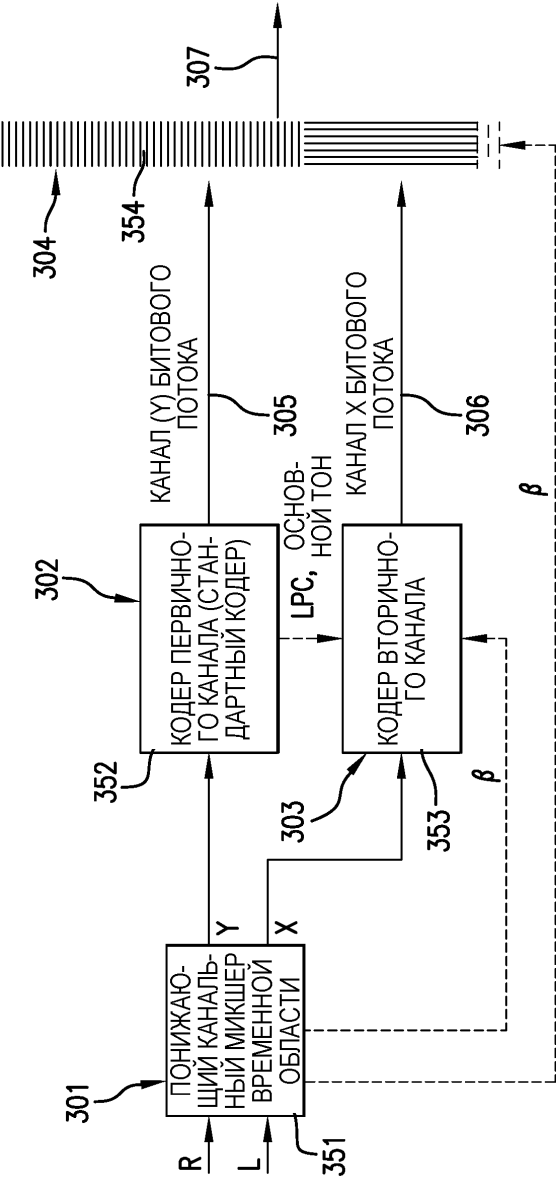
45



ФИГ. 1

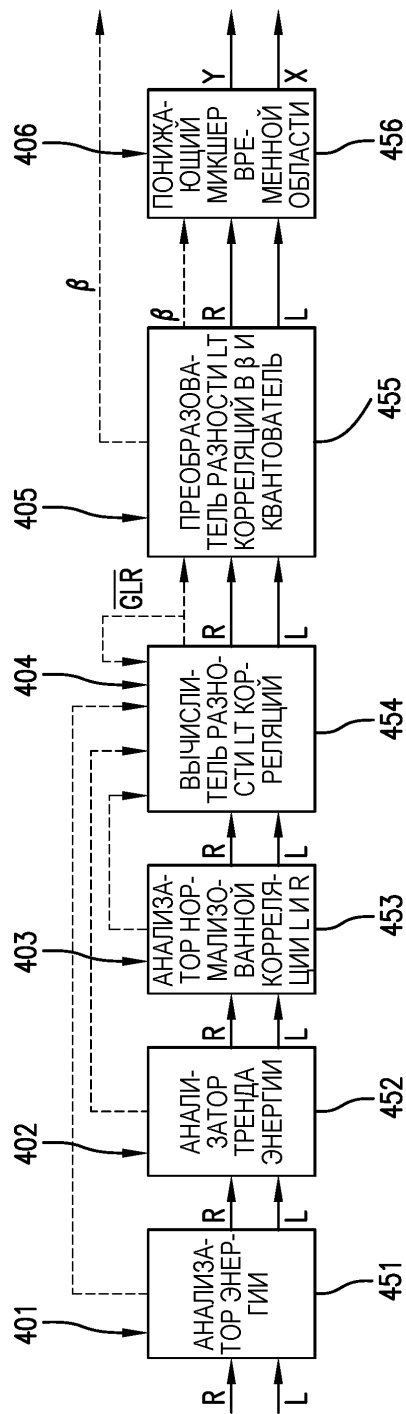


ФИГ. 2



ФИГ. 3

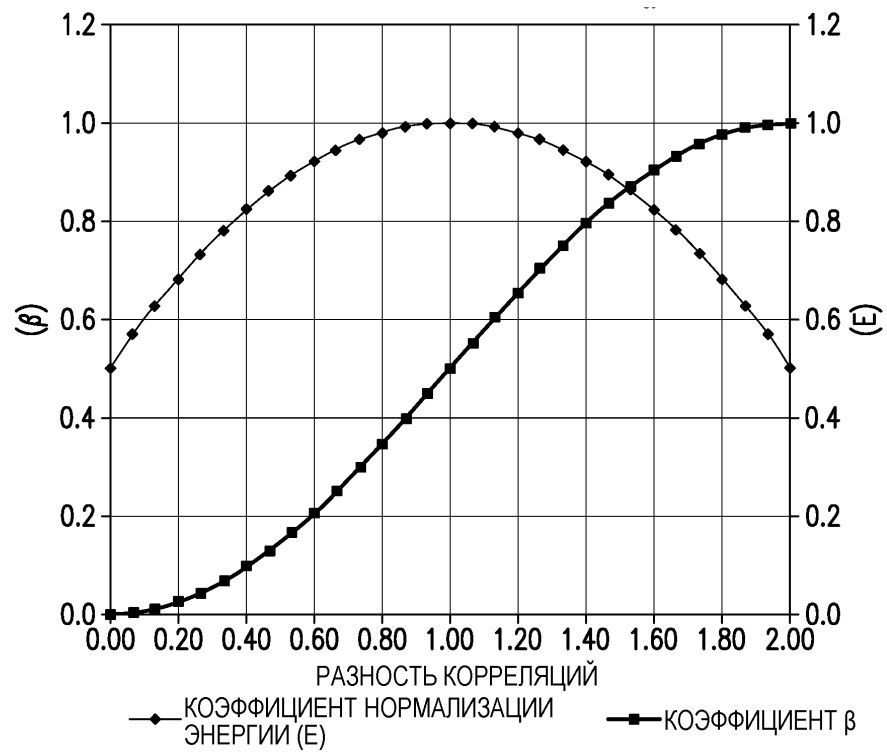
4/18



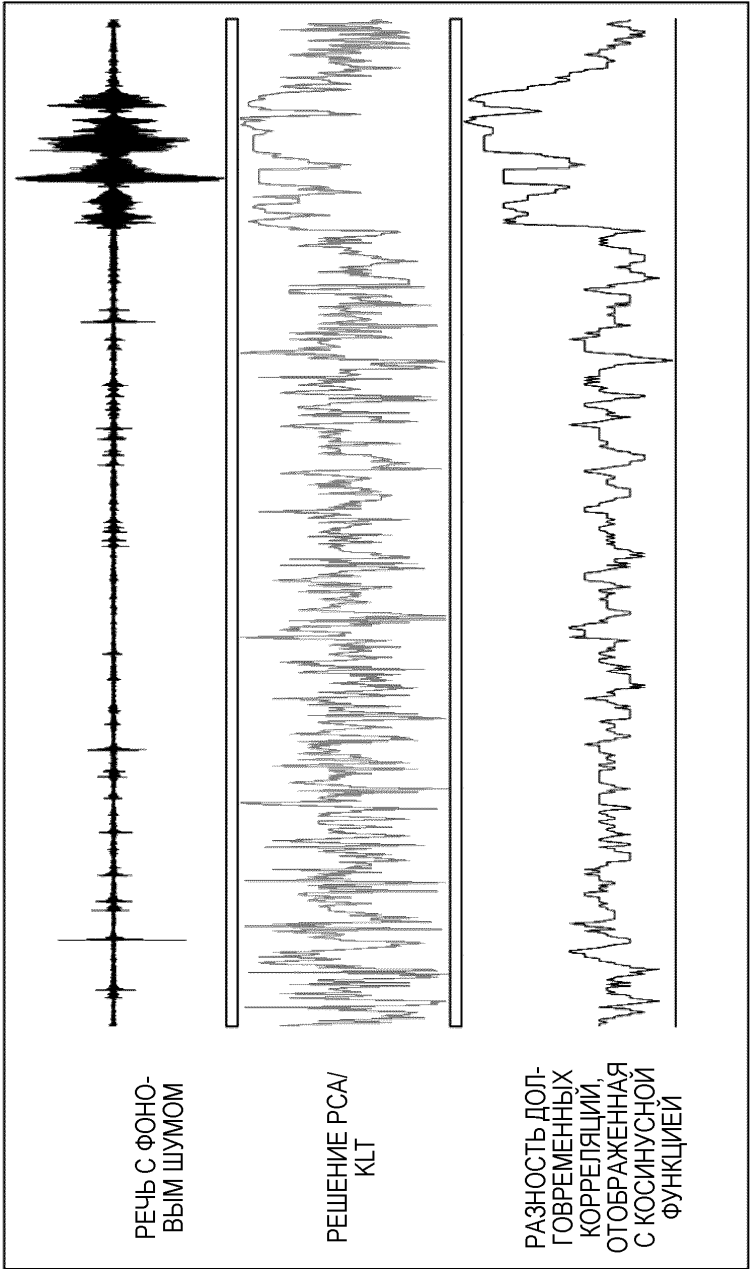
ФИГ. 4

5/18

РАЗНОСТЬ КОРРЕЛЯЦИЙ В КОСИНУСНОЙ ОБЛАСТИ И КОЭФФИЦИЕНТ НОРМАЛИЗАЦИИ ЭНЕРГИИ

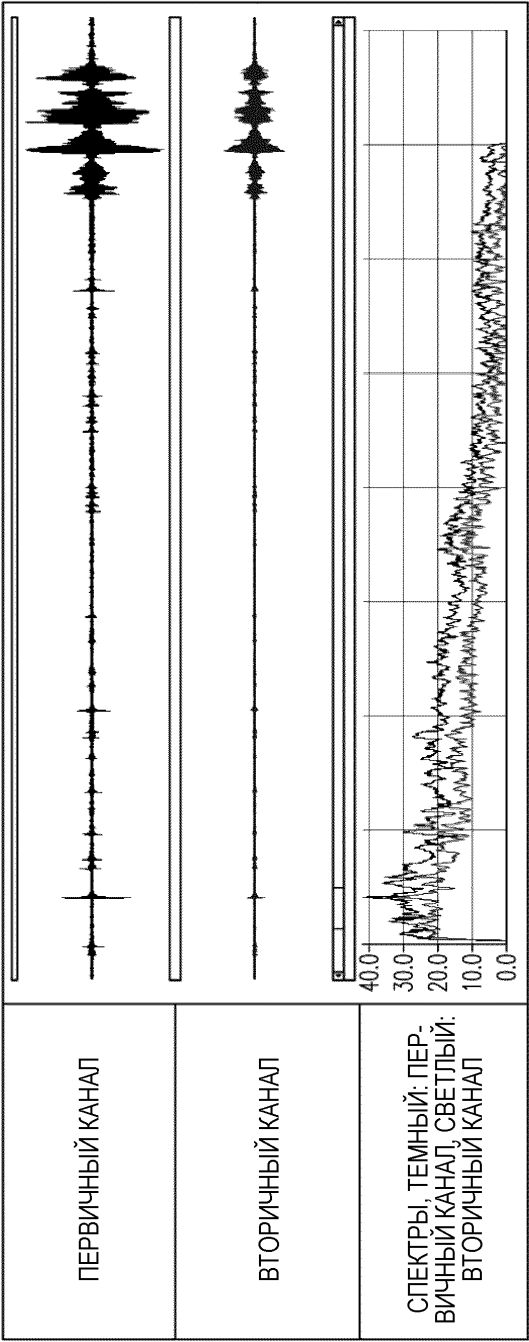


ФИГ. 5

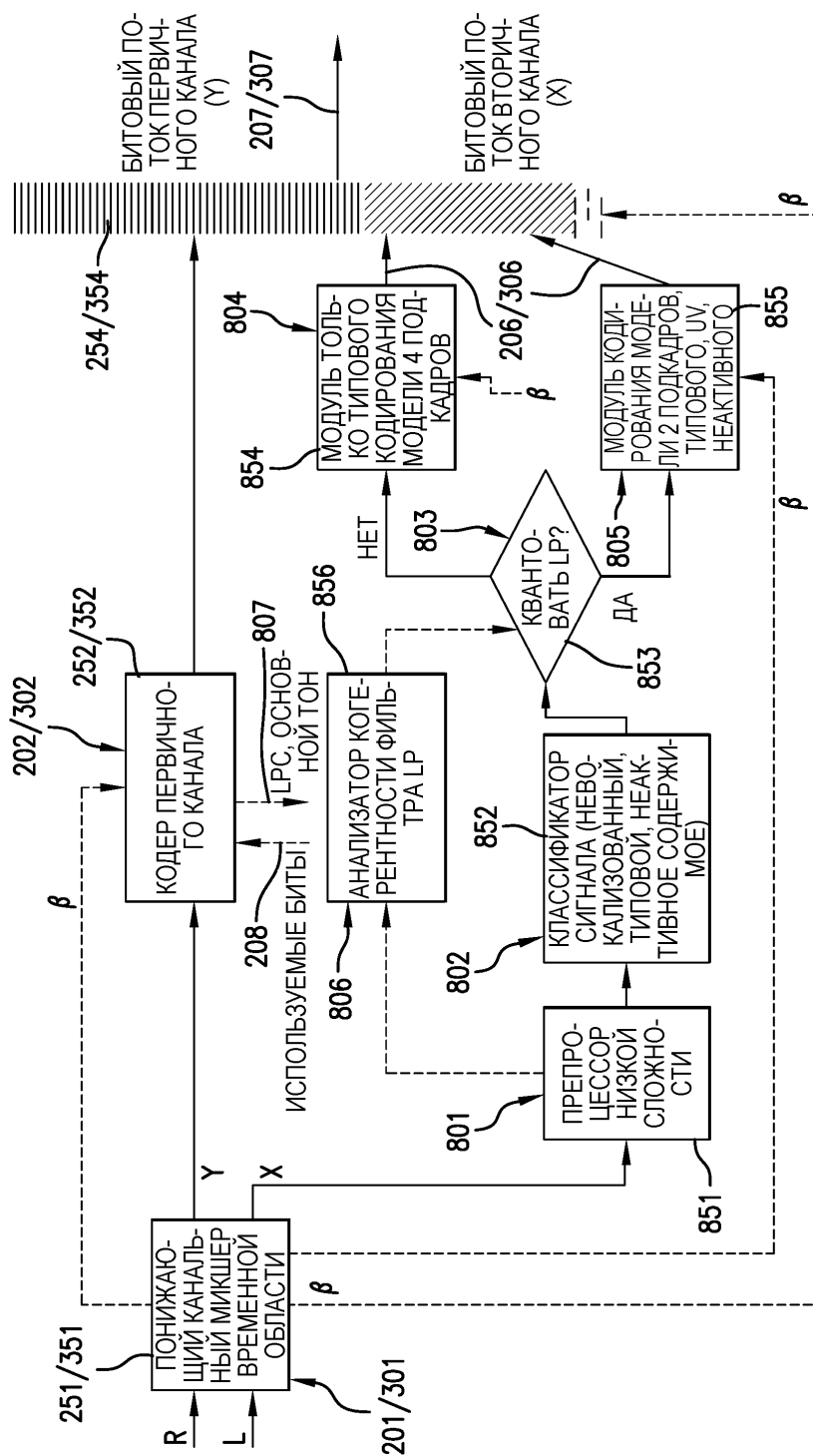


ФИГ. 6

7/18

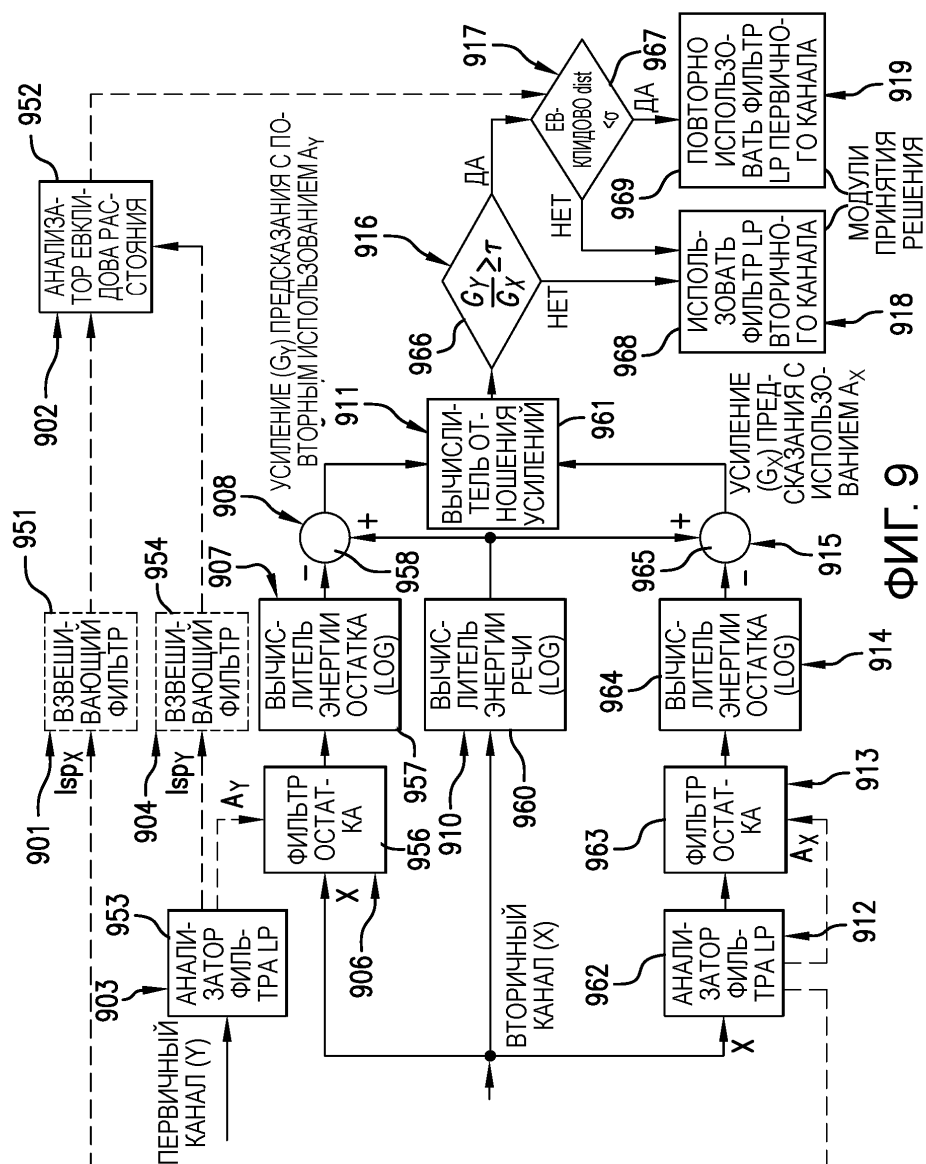


ФИГ. 7



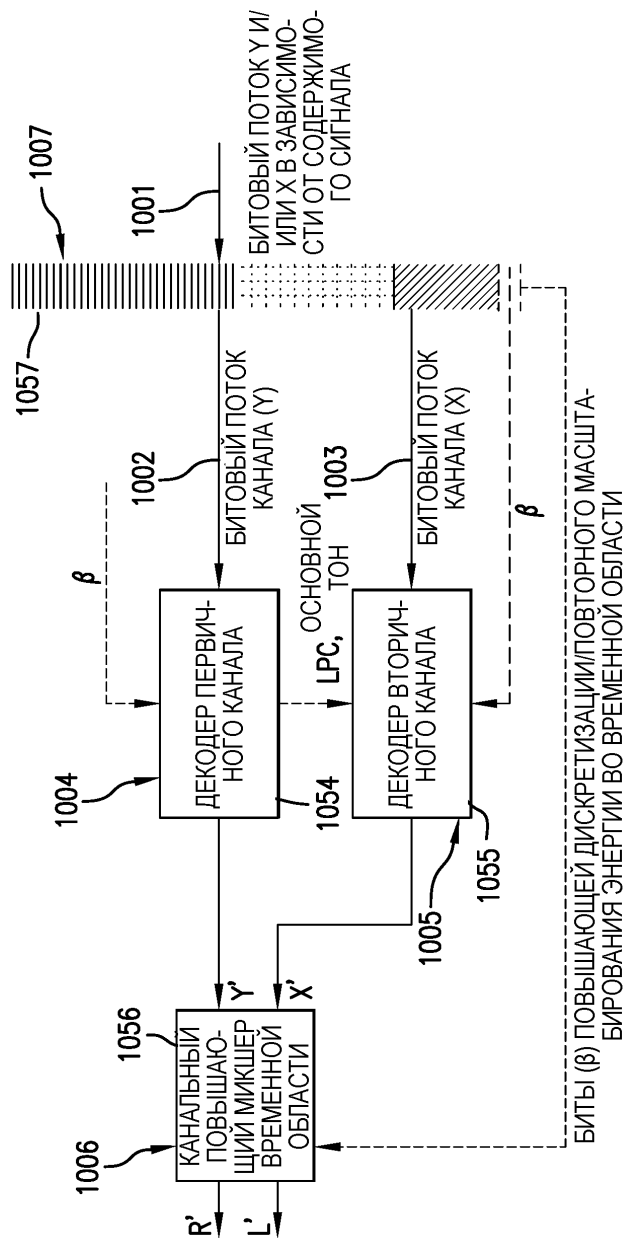
Φιν. 8

9/18

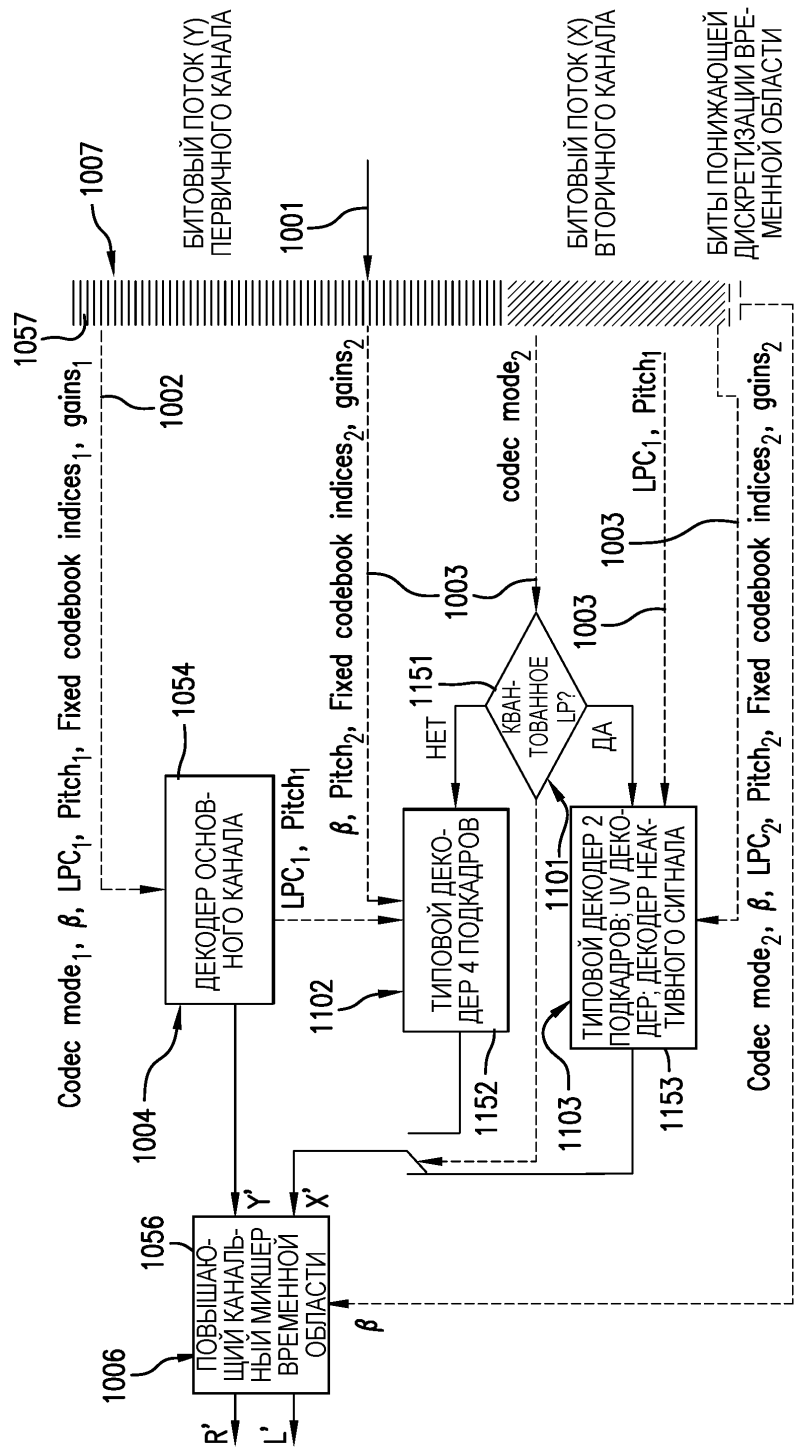


Φιν. 9

10/18

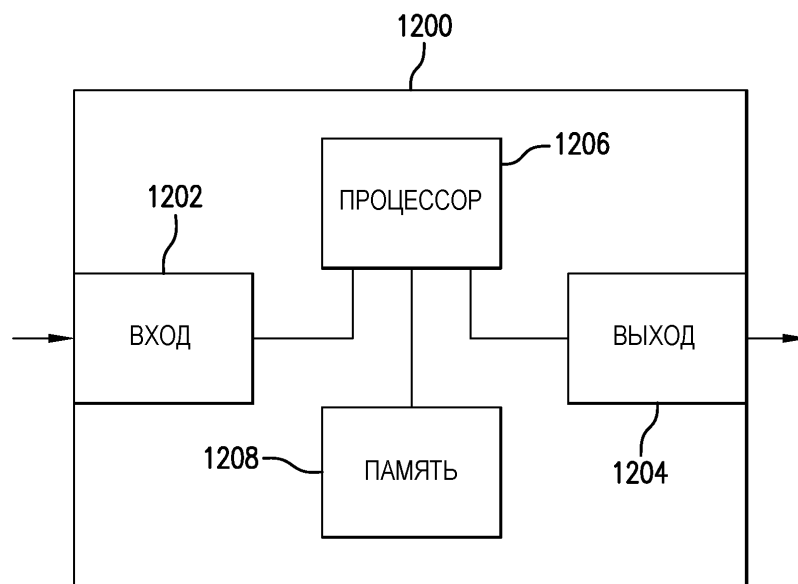


ФИГ. 10



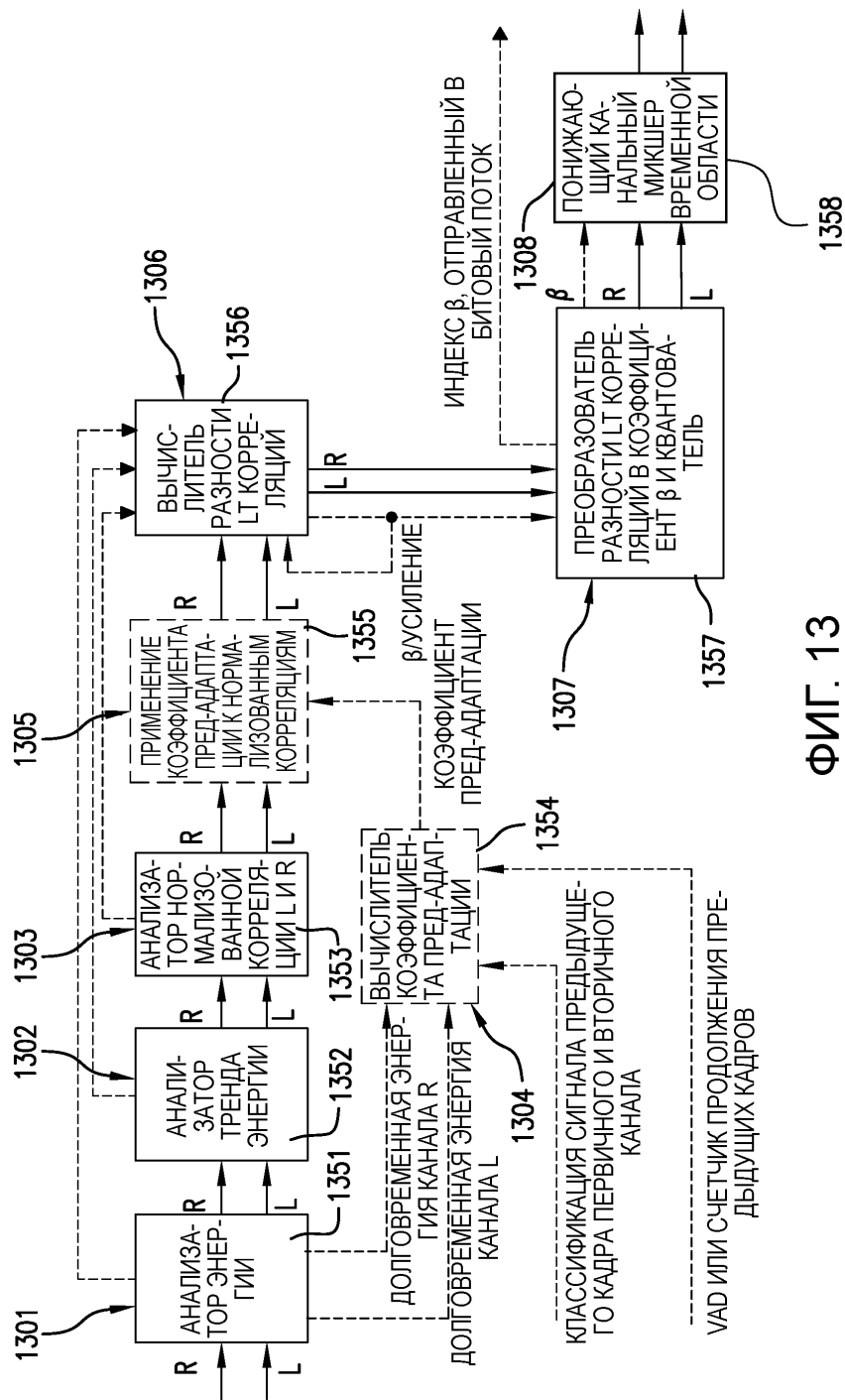
ФИГ. 11

12/18



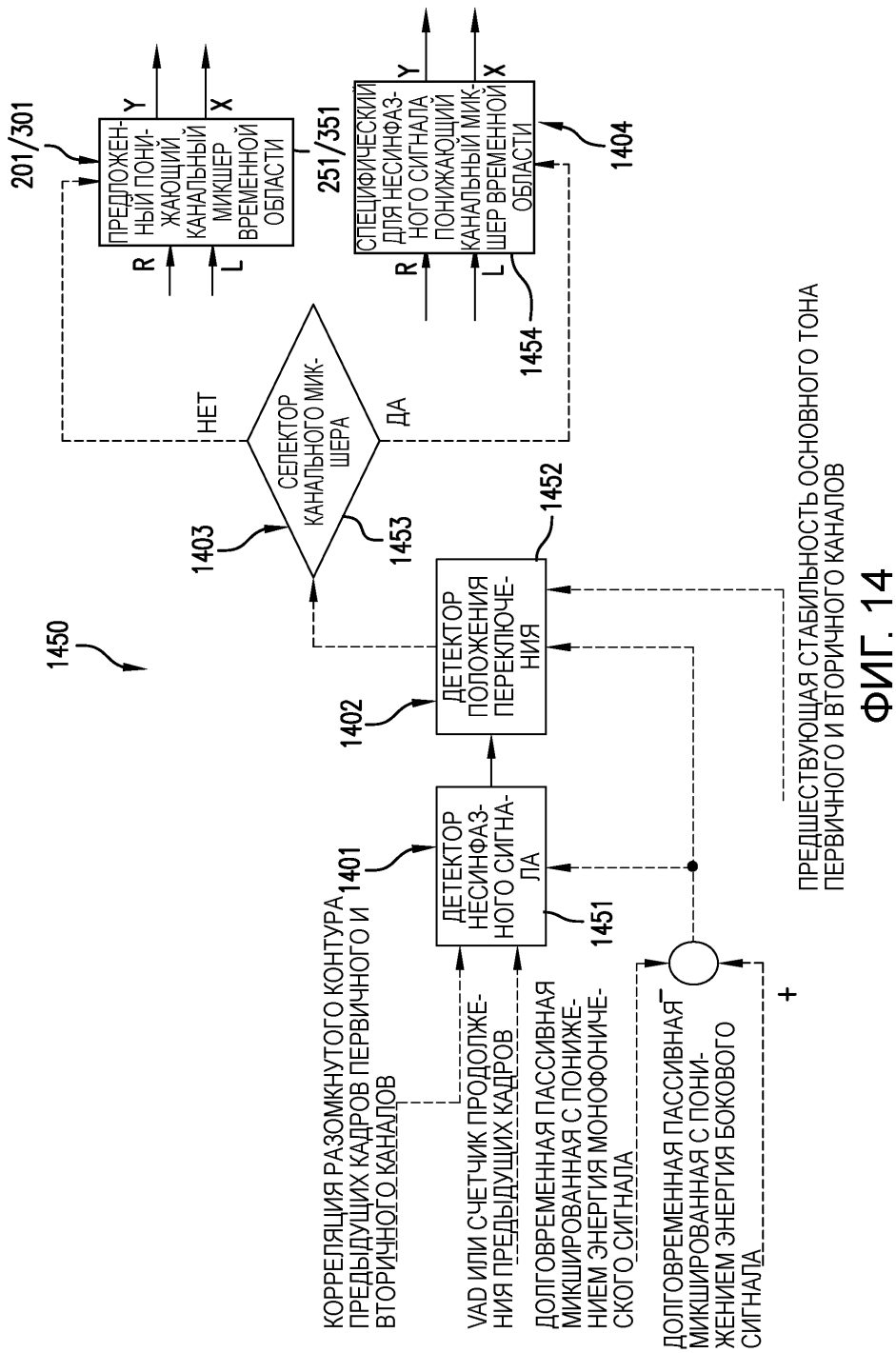
ФИГ. 12

13/18

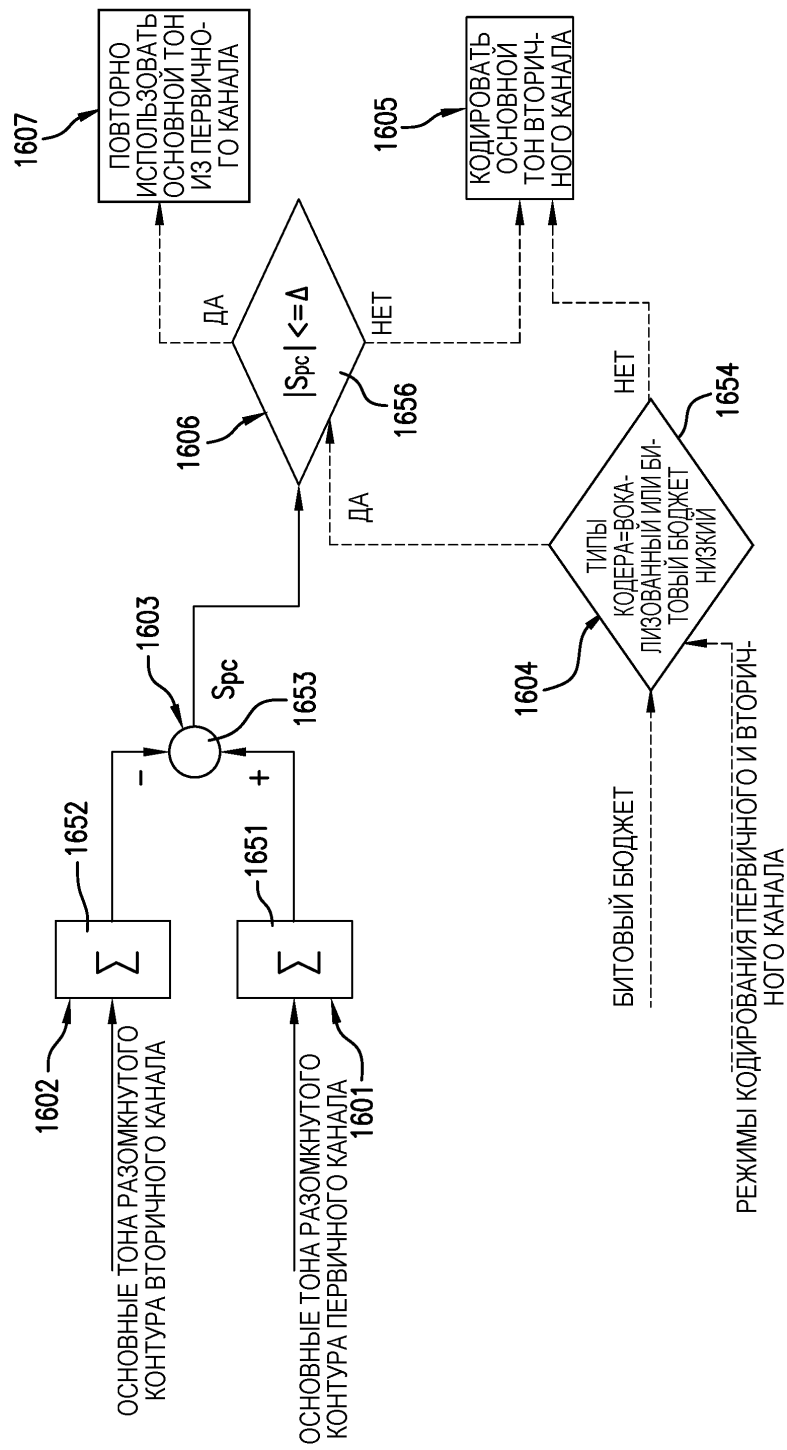


ФИГ. 13

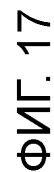
14/18

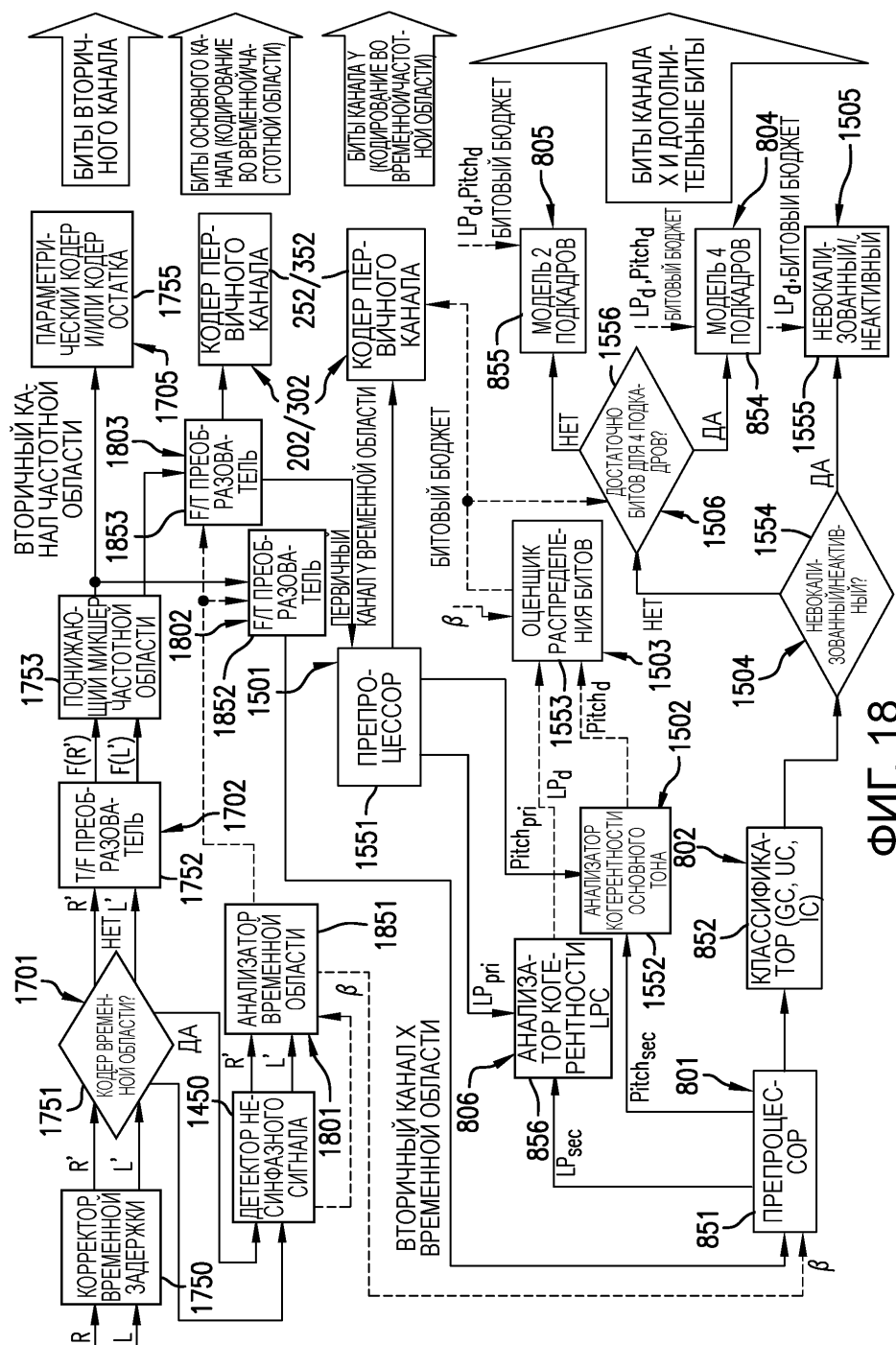


16/18



ФИГ. 16





Φιν. 18