



ФЕДЕРАЛЬНАЯ СЛУЖБА
ПО ИНТЕЛЛЕКТУАЛЬНОЙ СОБСТВЕННОСТИ

(12) ОПИСАНИЕ ИЗОБРЕТЕНИЯ К ПАТЕНТУ

(21)(22) Заявка: 2015147158, 04.04.2014

(24) Дата начала отсчета срока действия патента:
04.04.2014

Дата регистрации:
13.07.2017

Приоритет(ы):

(30) Конвенционный приоритет:
05.04.2013 US 61/809,019;
10.09.2013 US 61/875,959

(43) Дата публикации заявки: 17.05.2017 Бюл. № 14

(45) Опубликовано: 13.07.2017 Бюл. № 20

(85) Дата начала рассмотрения заявки РСТ на
национальной фазе: 05.11.2015

(86) Заявка РСТ:
EP 2014/056857 (04.04.2014)

(87) Публикация заявки РСТ:
WO 2014/161996 (09.10.2014)

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

(72) Автор(ы):

**ЧЕРЛИНГ, Кристофер (SE),
ПУРНХАГЕН, Хейко (SE),
ВИЛЛЕМОЕС, Ларс (SE)**

(73) Патентообладатель(и):

ДОЛБИ ИНТЕРНЭШНЛ АБ (NL)

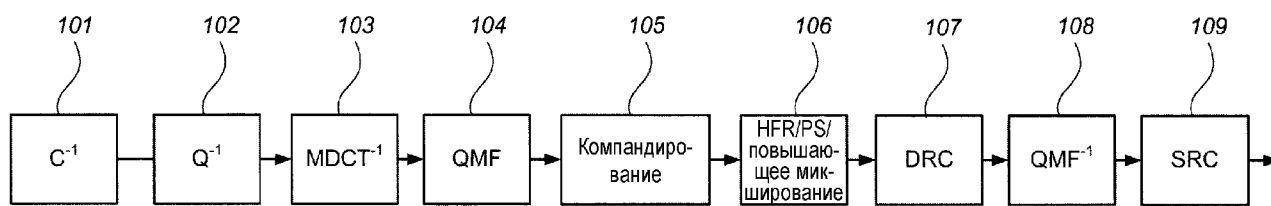
(56) Список документов, цитированных в отчете
о поиске: EP 2302624 A1, 30.03.2011. RU
2367033 C2, 10.09.2009. RU 2355046 C2,
20.05.2008. RU 2407073 C2, 20.12.2010. EP
2360683 A1, 24.08.2011. WO 2005/078076 A1,
25.08.2005. US 2010/0258542 A1, 21.10.2010. US
2011/0112829 A1, 12.05.2011.

(54) СИСТЕМА ОБРАБОТКИ АУДИО

(57) Реферат:

Изобретение относится к аудиокодированию и декодированию. Технический результат – восстановление дискретизированной звуковой волны в одноканальном, стереофоническом или многоканальном формате. Система обработки аудио содержит компонент предварительной обработки, который принимает квантованные спектральные компоненты и выполняет обратное квантование, получая в результате представление временной области для промежуточного сигнала, каскад обработки в частотной области, сконфигурированный для предоставления представления временной области для обработанного аудиосигнала, и преобразователь частоты дискретизации, предоставляющий

восстановленный аудиосигнал, дискретизированный с целевой частотой дискретизации. Соответствующие внутренние частоты дискретизации представления временной области для промежуточного аудиосигнала и представления временной области для обработанного аудиосигнала равны. В конкретных вариантах осуществления каскад обработки содержит каскад параметрического повышающего микширования, который работает в двух разных режимах и ассоциируется с каскадом задержки, который обеспечивает постоянную суммарную задержку. 3 н. и 12 з.п. ф-лы, 34 ил., 4 табл.



ФИГ. 1

R U 2 6 2 5 4 4 4 2 C 2

R U 2 6 2 5 4 4 4 2 C 2

(21)(22) Application: 2015147158, 04.04.2014

Registration date:
13.07.2017

Priority:

(30) Convention priority:
05.04.2013 US 61/809,019;
10.09.2013 US 61/875,959

(43) Application published: 17.05.2017 Bull. № 14

(45) Date of publication: 13.07.2017 Bull. № 20

(85) Commencement of national phase: **05.11.2015**

(86) PCT application:
EP 2014/056857 (04.04.2014)

(87) PCT publication:
WO 2014/161996 (09.10.2014)

Mail address:
129090, Moskva, ul. B. Spasskaya, 25, str. 3, OOO
"Yuridicheskaya firma Gorodisskij i Partnery"

(72) Inventor(s):

CHERLING, Kristofer (SE),
PURNKHAGEN, Khejko (SE),
VILLEMOES, Lars (SE)

(73) Proprietor(s):

DOLBY INTERNESHNL AB (NL)

(54) AUDIO PROCESSING SYSTEM

(57) Abstract:

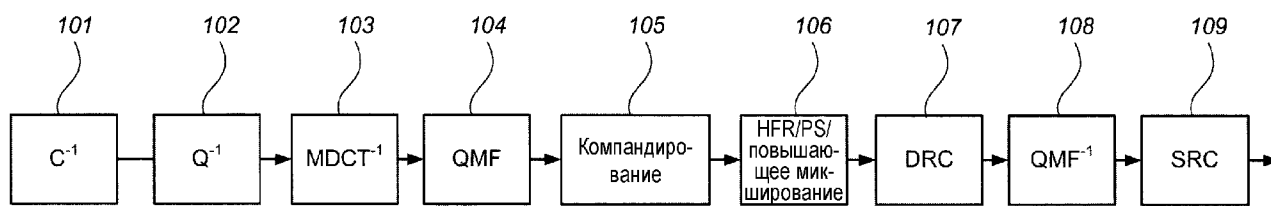
FIELD: physics.

SUBSTANCE: audio processing system comprises a pre-processing component, that receives the quantized spectral components and performs the inverse quantization, thereby obtaining a time domain representation for the intermediate signal, a processing stage in the frequency domain, configured to provide the time domain representation to the processed audio signal, and a sampling rate converter providing the reconstructed audio signal sampled with the target sampling frequency. Corresponding internal sampling

frequencies of the time domain representation for the intermediate audio signal and the time domain representation for the processed audio signal. In specific embodiments, the processing stage comprises a parametric upmixing stage, which operates in two different modes, and is associated with the delay stage, which provides a constant total delay.

EFFECT: sampled sound wave restoration in the single-channel, stereo or multi-channel format.

15 cl, 34 dwg, 4 tbl



ФИГ. 1

R U 2 6 2 5 4 4 4 4 4 2 C 2

R U 2 6 2 5 4 4 4 4 4 2 C 2

ПЕРЕКРЕСТНАЯ ССЫЛКА НА РОДСТВЕННЫЕ ЗАЯВКИ

Данная заявка испрашивает приоритет по предварительной патентной заявке США № 61/809,019, поданной 5 апреля 2013 г., и 61/875,959, поданной 10 сентября 2013 г., каждая из которых настоящим полностью включается в этот документ посредством

отсылки.

ОБЛАСТЬ ТЕХНИКИ

Данное раскрытие изобретения в целом относится к аудиокодированию и декодированию. Различные варианты осуществления предоставляют системы аудиокодирования и декодирования (называемые системами аудиокодеков), особенно

УРОВЕНЬ ТЕХНИКИ

Сложные технологические системы, включающие в себя системы аудиокодеков, обычно развиваются с нарастанием за длительный период времени и часто благодаря неслаженным усилиям в независимых научно-исследовательских коллективах. В результате такие системы могут включать в себя неудобные сочетания компонентов, которые представляют разные парадигмы проектирования и/или неодинаковые уровни технологического прогресса. Частое пожелание сохранить совместимость с унаследованным оборудованием накладывает дополнительное ограничение на проектировщиков и может привести к менее связной архитектуре системы. В системах параметрических многоканальных аудиокодеков обратная совместимость может включать в себя, в частности, предоставление кодированного формата, где сигнал понижающего микширования вернет оптимально звучащий выход при проигрывании в системе монофонического или стереофонического воспроизведения без возможностей обработки.

Доступные форматы аудиокодирования, представляющие уровень техники, включают в себя MPEG Surround, USAC и Высокоэффективное AAC v2. Они всесторонне описаны и проанализированы в литературе.

Было бы желательно предложить универсальную и к тому же архитектурно однородную систему аудиокодека с приемлемой производительностью, особенно для речевых сигналов.

КРАТКОЕ ОПИСАНИЕ ЧЕРТЕЖЕЙ

Далее будут подробно описываться варианты осуществления в рамках идеи изобретения со ссылкой на прилагаемые чертежи, на которых

фиг.1 - обобщенная блок-схема, показывающая общую структуру системы обработки аудио в соответствии с примерным вариантом осуществления;

фиг.2 показывает тракты обработки для двух разных режимов монофонического декодирования в системе обработки аудио;

фиг.3 показывает тракты обработки для двух разных режимов параметрического стереофонического декодирования, без включения дополнения после повышающего микширования кодированным по форме низкочастотным содержимым и с включением такого,

фиг.4 показывает тракт обработки для режима декодирования, в котором система обработки аудио обрабатывает полностью кодированный по форме стереофонический сигнал с дискретно кодированными каналами;

фиг.5 показывает тракт обработки для режима декодирования, в котором система обработки аудио предоставляет пятиканальный сигнал путем параметрического повышающего микширования трехканального сигнала понижающего микширования после применения копирования спектральных полос;

фиг.6 показывает структуру системы обработки аудио в соответствии с примерным вариантом осуществления, а также внутренние механизмы компонента в системе;

фиг.7 - обобщенная блок-схема системы декодирования в соответствии с примерным вариантом осуществления;

5 фиг.8 иллюстрирует первую часть системы декодирования на фиг.7;

фиг.9 иллюстрирует вторую часть системы декодирования на фиг.7;

фиг.10 иллюстрирует третью часть системы декодирования на фиг.7;

фиг.11 - обобщенная блок-схема системы декодирования в соответствии с примерным вариантом осуществления;

10 фиг.12 иллюстрирует третью часть системы декодирования из фиг.11; и

фиг.13 - обобщенная блок-схема системы декодирования в соответствии с примерным вариантом осуществления;

фиг.14 иллюстрирует первую часть системы декодирования на фиг.13;

фиг.15 иллюстрирует вторую часть системы декодирования на фиг.13;

15 фиг.16 иллюстрирует третью часть системы декодирования на фиг.13;

фиг.17 - обобщенная блок-схема системы кодирования в соответствии с первым примерным вариантом осуществления;

фиг.18 - обобщенная блок-схема системы кодирования в соответствии со вторым примерным вариантом осуществления;

20 фиг.19a показывает блок-схему примерного аудиокодера, предоставляющего поток битов с постоянной скоростью передачи битов;

фиг.19b показывает блок-схему примерного аудиокодера, предоставляющего поток битов с переменной скоростью передачи битов;

25 фиг.20 иллюстрирует формирование примерной огибающей на основе множества блоков коэффициентов преобразования;

фиг.21a иллюстрирует примерные огибающие у блоков коэффициентов преобразования;

фиг.21b иллюстрирует определение примерной интерполированной огибающей;

фиг.22 иллюстрирует примерные наборы квантователей;

30 фиг.23a показывает блок-схему примерного аудиодекодера;

фиг.23b показывает блок-схему примерного декодера огибающей в аудиодекодере из фиг.23a;

фиг.23c показывает блок-схему примерного блока предсказания субполосы в аудиодекодере из фиг.23a;

35 фиг.23d показывает блок-схему примерного декодера спектра в аудиодекодере из фиг.23a;

фиг.24a показывает блок-схему примерного набора допустимых квантователей;

фиг.24b показывает блок-схему примерного квантователя с добавлением псевдослучайного шума;

40 фиг.24c иллюстрирует примерный выбор квантователей на основе спектра блока коэффициентов преобразования;

фиг.25 иллюстрирует примерную схему для определения набора квантователей в кодере и в соответствующем декодере;

45 фиг.26 показывает блок-схему примерной схемы для декодирования энтропийно кодированных индексов квантования, которые определены с использованием квантователя с добавлением псевдослучайного шума; и

фиг.27 иллюстрирует примерный процесс распределения битов.

Все фигуры являются схематическими и показывают, как правило, только части,

которые необходимы, чтобы объяснить изобретение, тогда как другие части могут пропускаться или всего лишь предполагаться.

ПОДРОБНОЕ ОПИСАНИЕ

Система обработки аудио принимает аудиопоток битов, сегментированный на кадры, переносимые аудиоданные. Аудиоданные могут быть подготовлены путем дискретизации звуковой волны и преобразования полученных таким образом электронных временных выборок в спектральные коэффициенты, которые затем квантуют и кодируют в формат, подходящий для передачи или хранения. Система обработки аудио приспособлена для восстановления дискретизированной звуковой волны в одноканальном, стереофоническом или многоканальном формате. При использовании в данном документе аудиосигнал может относиться к чистому аудиосигналу либо к аудиочасти видеосигнала, аудиовизуального или мультимедийного сигнала.

Система обработки аудио, как правило, разделяется на компонент предварительной обработки, каскад обработки и преобразователь частоты дискретизации. Компонент предварительной обработки включает в себя: каскад деквантования, приспособленный для приема квантованных спектральных коэффициентов и для вывода первого представления частотной области для промежуточного сигнала; и каскад обратного преобразования для приема первого представления частотной области для промежуточного сигнала и синтеза на его основе представления временной области для промежуточного сигнала. Каскад обработки, который в некоторых вариантах осуществления можно полностью обходить, включает в себя: гребенку фильтров анализа для приема представления временной области для промежуточного сигнала и вывода второго представления частотной области для промежуточного сигнала; по меньшей мере один компонент обработки для приема упомянутого второго представления частотной области для промежуточного сигнала и вывода представления частотной области для обработанного аудиосигнала; и гребенку фильтров синтеза для приема представления частотной области для обработанного аудиосигнала и вывода представления временной области для обработанного аудиосигнала. Преобразователь частоты дискретизации в конечном счете конфигурируется для приема представления временной области для обработанного аудиосигнала и для вывода восстановленного аудиосигнала, дискретизированного с целевой частотой дискретизации.

В соответствии с примерным вариантом осуществления система обработки аудио имеет одночастотную архитектуру, в которой равны соответствующие внутренние частоты дискретизации представления временной области для промежуточного аудиосигнала и представления временной области для обработанного аудиосигнала.

В конкретных примерных вариантах осуществления, в которых каскад предварительной обработки содержит базовый кодировщик, а каскад обработки содержит каскад параметрического повышающего микширования, базовый кодировщик и каскад параметрического повышающего микширования работают с одинаковой частотой дискретизации. Дополнительно или в качестве альтернативы базовый кодировщик можно расширить для обработки большего диапазона длин преобразования, а преобразователь частоты дискретизации можно сконфигурировать для соответствия стандартным частотам видеокадров, чтобы сделать возможным декодирование синхронных с видео аудиокладов. Это будет подробнее описываться ниже в разделе "Кодирование в аудиорежиме".

В еще одних конкретных примерных вариантах осуществления компонент предварительной обработки работает в аудиорежиме и речевом режиме, отличном от аудиорежима. Поскольку речевой режим специально приспособлен для речевого

содержимого, такие сигналы могут проигрываться точнее. В аудиорежиме компонент предварительной обработки может работать аналогично тому, что раскрывается на фиг.6 и в связанных разделах данного описания. В речевом режиме компонент предварительной обработки может работать так, как подробно обсуждается ниже в

5 разделе "Кодирование в речевом режиме".

Вообще говоря, в примерных вариантах осуществления речевой режим отличается от аудиорежима компонента предварительной обработки в том, что каскад обратного преобразования работает с меньшей длиной кадра (или размером преобразования).

Выявлено, что сокращенная длина кадра эффективнее захватывает речевое содержимое.

10 В некоторых примерных вариантах осуществления длина кадра является переменной в аудиорежиме и в видеорежиме; ее можно, например, периодически уменьшать для захвата переходов в сигнале. В таких обстоятельствах изменение режима с аудиорежима на речевой режим - при прочих равных условиях - будет подразумевать сокращение длины кадра в каскаде обратного преобразования. Иначе говоря, такое изменение

15 режима с аудиорежима на речевой режим будет подразумевать сокращение максимальной длины кадра (среди выбираемых длин кадров в каждом из аудиорежима и речевого режима). В частности, длина кадра в речевом режиме может быть постоянной долей (например, 1/8) текущей длины кадра в аудиорежиме.

В примерном варианте осуществления обходная линия, параллельная каскаду

20 обработки, позволяет обходить каскад обработки в режимах декодирования, где не нужна обработка в частотной области. Это может быть применимым, когда система декодирует дискретно кодированные стереофонические или многоканальные сигналы, в частности, сигналы, где кодирован по форме полный спектральный диапазон (в соответствии с чем может не требоваться копирование спектральных полос). Чтобы

25 избежать временных сдвигов в случаях, когда обходная линия включается в тракт обработки или исключается из него, обходная линия предпочтительно может содержать каскад задержки, соответствующий задержке (или алгоритмической задержке) каскада обработки в текущем режиме. В вариантах осуществления, в которых каскад обработки выполнен с возможностью иметь постоянную (алгоритмическую) задержку независимо

30 от его текущего режима работы, каскад задержки на обходной линии может вызывать постоянную, заранее установленную задержку; в противном случае каскад задержки на обходной линии предпочтительно является адаптивным и меняется в соответствии с текущим режимом работы каскада обработки.

В примерном варианте осуществления каскад параметрического повышающего

35 микширования работает в режиме, где принимает 3-канальный сигнал понижающего микширования и возвращает 5-канальный сигнал. При желании компонент копирования спектральных полос может располагаться раньше каскада параметрического повышающего микширования. В конфигурации каналов воспроизведения с тремя передними каналами (например, L, R, C) и двумя каналами окружения (например, Ls,

40 Rs), где у кодированного сигнала доминирует передний канал, этот примерный вариант осуществления может добиться более эффективного кодирования. В действительности доступная полоса пропускания у аудиопотока битов в основном расходуется на попытку кодировать по форме как можно больше из трех передних каналов. Устройство кодирования, готовящее аудиопоток битов, декодируемый системой обработки аудио,

45 может адаптивно выбирать декодирование в этом режиме путем измерения свойств кодируемого аудиосигнала. Примерный вариант осуществления процедуры повышающего микширования, состоящей в повышающем микшировании одного канала понижающего микширования в два канала, и соответствующей процедуры понижающего

микширования обсуждается ниже под заголовком "Стерефоническое кодирование".

В дальнейшем развитии предыдущего примерного варианта осуществления два из трех каналов в сигнале понижающего микширования соответствуют совместно кодированным каналам в аудиопотоке битов. Такое совместное кодирование может повлечь за собой, например, выделение масштаба одного канала по сравнению с другим каналом. Аналогичный подход реализован в интенсивном стерефоническом кодировании AAC, в котором два канала могут кодироваться как элемент канальной пары. Экспериментами по прослушиванию доказано, что на заданной скорости передачи битов воспринимаемое качество восстановленного аудиосигнала повышается, когда некоторые каналы сигнала понижающего микширования кодируются совместно.

В примерном варианте осуществления система обработки аудио дополнительно содержит модуль копирования спектральных полос. Модуль копирования спектральных полос (или каскад высокочастотного восстановления) подробнее обсуждается ниже под заголовком "Стерефоническое кодирование". Модуль копирования спектральных полос предпочтительно активен, когда каскад параметрического повышающего микширования выполняет операцию повышающего микширования, то есть когда он возвращает сигнал с большим количеством каналов, нежели сигнал, который он принимает. Однако, когда каскад параметрического повышающего микширования действует как транзитный компонент, модулем копирования спектральных полос можно управлять независимо от конкретного текущего режима у каскада параметрического повышающего микширования; иначе говоря, в режимах непараметрического декодирования функциональные возможности копирования спектральных полос необязательны.

В примерном варианте осуществления по меньшей мере один компонент обработки дополнительно включает в себя каскад кодирования по форме, который подробнее описывается ниже в разделе "Многоканальное кодирование".

В примерном варианте осуществления система обработки аудио работает для предоставления сигнала понижающего микширования, подходящего для унаследованного воспроизводящего оборудования. Точнее говоря, стерефонический сигнал понижающего микширования получается путем добавления синфазного содержимого канала окружения в первый канал в сигнал понижающего микширования и путем добавления сдвинутого по фазе (например, на 90 градусов) содержимого канала окружения во второй канал. Это позволяет воспроизводящему оборудованию вывести содержимое канала окружения путем объединенной операции обратного сдвига по фазе и вычитания. Сигнал понижающего микширования может быть допустимым для воспроизводящего оборудования, сконфигурированного для приема общего левого/общего правого сигнала понижающего микширования. Предпочтительно, чтобы функциональные возможности сдвига по фазе не были настройкой по умолчанию в системе обработки аудио, а могли отключаться, когда система обработки аудио готовит сигнал понижающего микширования, не предназначенный для воспроизводящего оборудования этого типа. В действительности известны отдельные типы содержимого, которые плохо воспроизводятся со сдвинутыми по фазе окружающими сигналами; в частности, звук, записанный из источника с ограниченным пространственным объемом, который затем панорамируется между левым передним и левым окружающим сигналом, как и предполагается, не будет восприниматься как расположенный между соответствующими левым передним и левым окружающими динамиками, но не будет ассоциирован с четким пространственным расположением в соответствии с мнением многих слушателей. Этого артефакта можно избежать путем реализации фазового

сдвига канала окружения в виде необязательных, нестандартных функциональных возможностей.

В примерном варианте осуществления компонент предварительной обработки содержит блок предсказания, декодер спектра, узел добавления и узел обратного
5 выравнивания. Эти элементы, которые повышают производительность системы, когда она обрабатывает сигналы речевого типа, будут подробнее описываться ниже под заголовком "Кодирование в речевом режиме".

В примерном варианте осуществления система обработки аудио дополнительно содержит декодер Lfe (низкочастотных эффектов) для подготовки по меньшей мере
10 одного дополнительного канала на основе информации в аудиопотоке битов. Предпочтительно, чтобы декодер Lfe предоставлял канал низкочастотных эффектов, который кодируется по форме, отдельно от других каналов, переносимых аудиопотоком битов. Если дополнительный канал кодируется дискретно с другими каналами восстановленного аудиосигнала, то соответствующий тракт обработки может не
15 зависеть от остальной части системы обработки аудио. Подразумевается, что каждый дополнительный канал добавляется к общему количеству каналов в восстановленном аудиосигнале; например, в варианте использования, где каскад параметрического повышающего микширования - если предоставляется - работает в режиме $N = 5$, и где имеется один дополнительный канал, общее количество каналов в восстановленном
20 аудиосигнале будет равно $N + 1 = 6$.

Дополнительные примерные варианты осуществления предоставляют способ, включающий в себя этапы, соответствующие операциям, выполняемым вышеупомянутой
системой обработки аудио, когда она используется, и компьютерный программный продукт для побуждения программируемого компьютера выполнить такой способ.

Идея изобретения дополнительно относится к системе обработки аудио кодирующего
25 типа для кодирования аудиосигнала в аудиопоток битов, имеющий подходящий формат для декодирования в описанной выше системе обработки аудио (декодирующего типа). Первая идея изобретения дополнительно включает в себя способы кодирования и компьютерные программные продукты для подготовки аудиопотока битов.

Фиг.1 показывает систему 100 обработки аудио в соответствии с примерным
30 вариантом осуществления. Базовый декодер 101 принимает аудиопоток битов и выводит, по меньшей мере, квантованные спектральные коэффициенты, которые поступают в компонент предварительной обработки, содержащий каскад 102 деквантования и каскад 103 обратного преобразования. Компонент предварительной обработки в некоторых
35 примерных вариантах осуществления может быть двухрежимным. В тех вариантах осуществления им можно выборочно управлять в универсальном аудиорежиме и в специальном аудиорежиме (например, в речевом режиме). После компонента предварительной обработки каскад обработки ограничивается гребенкой 104 фильтров анализа на входе и гребенкой 108 фильтров синтеза на выходе. Компоненты,
40 размещенные между гребенкой 104 фильтров анализа и гребенкой 108 фильтров синтеза, выполняют обработку в частотной области. В варианте осуществления первой идеи, показанном на фиг.1, эти компоненты включают в себя:

- компонент 105 компандирования;
- объединенный компонент 106 для высокочастотного восстановления,
- 45 параметрического стерео и повышающего микширования; и
- компонент 107 динамического управления диапазоном.

Компонент 106, например, может выполнять повышающее микширование, которое описано ниже в разделе "Стереофоническое кодирование" настоящего описания.

После каскада обработки система 100 обработки аудио дополнительно содержит преобразователь 109 частоты дискретизации, сконфигурированный для предоставления восстановленного аудиосигнала, дискретизированного с целевой частотой дискретизации.

На выходе система 100 при желании может включать в себя компонент ограничения сигнала (не показан), отвечающий за выполнение условия без срезания.

Кроме того, система 100 при желании может содержать параллельный тракт обработки для предоставления одного или нескольких дополнительных каналов (например, канала низкочастотных эффектов). Параллельный тракт обработки можно реализовать в виде декодера Lfe (не показан ни на какой из фиг.1 и 3-11), который принимает аудиопотоки битов или их часть и который выполнен с возможностью вставки подготовленного таким образом дополнительного канала (каналов) в восстановленный аудиосигнал; точка вставки может находиться непосредственно перед преобразователем 109 частоты дискретизации.

Фиг.2 иллюстрирует два режима монофонического декодирования в системе обработки аудио, показанной на фиг.1, с соответствующим обозначением. Точнее говоря, фиг.2 показывает те компоненты системы, которые активны в течение декодирования и которые образуют тракт обработки для подготовки восстановленного (монофонического) аудиосигнала на основе аудиопотока битов. Отметим, что тракты обработки на фиг.2 дополнительно включают в себя завершающий компонент ограничения сигнала ("Lim"), выполненный с возможностью уменьшения масштаба значений сигналов для выполнения условия без срезания. Верхний режим декодирования на фиг.2 использует высокочастотное восстановление, тогда как нижний режим декодирования на фиг.2 декодирует полностью кодированный по форме канал. Поэтому в нижнем режиме декодирования компонент высокочастотного восстановления ("HFR") заменен каскадом задержки ("Задержка"), вызывающим задержку, равную алгоритмической задержке компонента HFR.

Как предлагает нижняя часть фиг.2, дополнительно можно полностью обойти каскад обработки ("QMF", "Задержка", "DRC", "QMF⁻¹"); это может применяться, когда над сигналом не выполняется никакая обработка по динамическому управлению диапазоном (DRC). Обход каскада обработки устраняет любое возможное ухудшение сигнала из-за анализа QMF с последующим синтезом QMF, который может содержать несовершенное восстановление. Обходная линия включает в себя каскад второй линии задержки, сконфигурированный для задержки сигнала на величину, равную общей (алгоритмической) задержке каскада обработки.

Фиг.3 иллюстрирует два режима параметрического стереофонического декодирования. В обоих режимах стереофонические каналы получаются путем применения высокочастотного восстановления к первому каналу, которое дает его декоррелированную версию с использованием декоррелятора ("D"), а затем образования линейной комбинации обеих версий для получения стереофонического сигнала. Линейная комбинация вычисляется каскадом повышающего микширования ("Повышающее микширование"), размещенным раньше каскада DRC. В одном из режимов - который показан в нижней части чертежа - аудиопоток битов дополнительно переносит кодированное по форме низкочастотное содержимое для обоих каналов (область, заштрихованная "\ \ "). Подробности реализации последнего режима описываются с помощью фиг.7-10 и соответствующих разделов настоящего описания.

Фиг.4 иллюстрирует режим декодирования, в котором система обработки аудио обрабатывает полностью кодированный по форме стереофонический сигнал с дискретно

кодированными каналами. Это высокоскоростной стереофонический режим. Если обработка DRC не считается необходимой, то каскад обработки можно полностью обойти, используя две обходные линии с соответствующими каскадами задержки, показанными на фиг.4. Каскады задержки предпочтительно вызывают задержку, равную задержке у каскада обработки в других режимах декодирования, чтобы переключение режима могло происходить непрерывно относительно содержимого сигнала.

Фиг.5 иллюстрирует режим декодирования, в котором система обработки аудио предоставляет пятиканальный сигнал путем параметрического повышающего микширования трехканального сигнала понижающего микширования после применения копирования спектральных полос. Как уже упоминалось, выгодно кодировать два канала (область, заштрихованная "///") совместно (например, в виде элемента канальной пары), и система обработки аудио предпочтительно проектируется для обработки потока битов с этим свойством. С этой целью система обработки аудио содержит два приемных участка, причем нижний конфигурируется для декодирования элемента канальной пары, а верхний - для декодирования оставшегося канала (область, заштрихованная "\\"). После высокочастотного восстановления в области QMF каждый канал в канальной паре декоррелируется отдельно, после чего первый каскад повышающего микширования образует первую линейную комбинацию первого канала и его декоррелированной версии, а второй каскад повышающего микширования образует вторую линейную комбинацию второго канала и его декоррелированной версии. Подробности реализации этой обработки описываются с помощью фиг.7-10 и соответствующих разделов настоящего описания. Все пять каналов затем подвергаются обработке DRC перед синтезом QMF.

Кодирование в аудиорежиме

Фиг.6 - обобщенная блок-схема системы 100 обработки аудио, принимающей кодированный аудиопоток Р битов, с восстановленным аудиосигналом, показанным в виде пары стереофонических основоположных сигналов L, R на фиг.6, в качестве окончательного результата. В этом примере допустим, что поток Р битов содержит квантованные двухканальные аудиоданные, кодированные с преобразованием. Система 100 обработки аудио может принимать аудиопоток Р битов из сети связи, от беспроводного приемника или из запоминающего устройства (не показано). Выход системы 100 может поступать в громкоговорители для воспроизведения или может перекодироваться в таком же или другом формате для дальнейшей передачи по сети связи либо беспроводной линии связи или для сохранения в запоминающем устройстве.

Система 100 обработки аудио содержит декодер 108 для декодирования потока Р битов на квантованные спектральные коэффициенты и управляющие данные. Компонент 110 предварительной обработки, структура которого подробнее будет обсуждаться ниже, деквантует эти спектральные коэффициенты и выдает представление временной области для промежуточного аудиосигнала, обрабатываемое каскадом 120 обработки. Промежуточный аудиосигнал преобразуется гребенками 122_L, 122_R фильтров анализа во вторую частотную область, отличную от области, ассоциированной с ранее упомянутым кодирующим преобразованием; второе представление частотной области может быть представлением квадратурного зеркального фильтра (QMF), и в этом случае гребенки 122_L, 122_R фильтров анализа могут предоставляться в виде гребенок фильтров QMF. После гребенок 122_L, 122_R фильтров анализа модуль 124 копирования спектральных полос (SBR), отвечающий за высокочастотное восстановление, и модуль 126 динамического управления диапазоном (DRC) обрабатывают второе представление

частотной области для промежуточного аудиосигнала. После них гребенки 128_L , 128_R фильтров синтеза создают представление временной области для обработанного таким образом аудиосигнала. Как поймет специалист после изучения данного раскрытия изобретения, ни модуль 124 копирования спектральных полос, ни модуль 126 динамического управления диапазоном не являются необходимыми элементами изобретения; наоборот, система обработки аудио в соответствии с другим примерным вариантом осуществления может включать в себя дополнительные или альтернативные модули в каскаде 120 обработки. После каскада 120 обработки преобразователь 130 частоты дискретизации работает для регулирования частоты дискретизации обработанного аудиосигнала до нужной частоты дискретизации аудио, например 44,1 кГц или 48 кГц, для которой спроектировано предназначенное воспроизводящее оборудование (не показано). В данной области техники хорошо известно, как спроектировать преобразователь 130 частоты дискретизации с низким количеством артефактов на выходе. Преобразователь 130 частоты дискретизации можно отключать в моменты, где не нужно преобразование частоты дискретизации - то есть там, где каскад 120 обработки выдает обработанный аудиосигнал, который уже имеет целевую частоту дискретизации. Необязательный модуль 140 ограничения сигнала, размещенный после преобразователя 130 частоты дискретизации, конфигурируется для ограничения значений основополосного сигнала при необходимости в соответствии с условием без срезания, которое опять может выбираться в связи с конкретным предназначенным воспроизводящим оборудованием.

Как показано в нижней части фиг.6, компонент 110 предварительной обработки содержит каскад 114 деквантования, которым можно управлять в одном из нескольких режимов с разными размерами блоков, и каскад 118_L , 118_R обратного преобразования, который также может работать с разными размерами блоков. Предпочтительно, чтобы изменения режима у каскада 114 деквантования и каскада 118_L , 118_R обратного преобразования были синхронными, так что размер блока все время совпадает. Перед этими компонентами компонент 110 предварительной обработки содержит демультимплексор 112 для отделения квантованных спектральных коэффициентов от управляющих данных; обычно он перенаправляет управляющие данные в каскад 118_L , 118_R обратного преобразования и перенаправляет квантованные спектральные коэффициенты (и управляющие данные, при желании) в каскад 114 деквантования. Каскад 114 деквантования выполняет отображение из одного кадра индексов квантования (обычно представленных целыми числами) в один кадр спектральных коэффициентов (обычно представленных числами с плавающей запятой). Каждый индекс квантования ассоциируется с уровнем квантования (или точкой восстановления). Предполагая, что аудиопоток битов подготовлен с использованием неравномерного квантования, как обсуждалось выше, эта ассоциация не уникальна, пока не задается, к какой полосе частот относится индекс квантования. Иначе говоря, процесс деквантования может придерживаться разной кодовой книги для каждой полосы частот, и набор кодовых книг может меняться в зависимости от длины кадра и/или скорости передачи битов. На фиг.6 это проиллюстрировано схематически, где вертикальная ось обозначает частоту, а горизонтальная ось обозначает выделенное количество битов кодирования на единичную частоту. Отметим, что полосы частот обычно шире для более высоких частот и оканчиваются на половине внутренней частоты f_1 дискретизации. Внутреннюю частоту дискретизации можно отобразить в отличную в числовом отношении физическую частоту дискретизации в результате передискретизации в

преобразователе 130 частоты дискретизации; например, повышающая дискретизация на 4,3% отобразит $f_i = 46,034$ кГц в приблизительную физическую частоту 48 кГц и увеличит границы полосы низких частот с таким же коэффициентом. Как дополнительно предлагает фиг.6, кодер, готовящий аудиопоток битов, обычно выделяет разные количества битов кодирования разным полосам частот в соответствии со сложностью кодированного сигнала и предполагаемыми колебаниями чувствительности слуха человека.

В таблице 1 приводятся количественные данные, характеризующие режимы работы системы 100 обработки аудио и, в частности, компонента 110 предварительной обработки.

Таблица 1 Примерные режимы а-м работы системы обработки аудио									
Режим	Частота кадров	Длительность кадра	Длина кадра в компоненте предварительной обработки	Ширина элемента разрешения в компоненте предварительной обработки	Внутренняя частота дискретизации	Гребенка фильтров анализа	Ширина полосы частот анализа	Коэффициент SRC	Внешняя частота дискретизации
	[Гц]	[мс]	[выборки]	[Гц]	[кГц]	[полосы]	[Гц]		[кГц]
A	23,976	41,708	1920	11,988	46,034	64	359,640	0,9590	48,000
B	24,000	41,667	1920	12,000	46,080	64	360,000	0,9600	48,000
C	24,975	40,040	1920	12,488	47,952	64	374,625	0,9990	48,000
D	25,000	40,000	1920	12,500	48,000	64	375,000	1,0000	48,000
E	29,970	33,367	1536	14,985	46,034	64	359,640	0,9590	48,000
F	30,000	33,333	1536	15,000	46,080	64	360,000	0,9600	48,000
G	47,952	20,854	960	23,976	46,034	64	359,640	0,9590	48,000
H	48,000	20,833	960	24,000	46,080	64	360,000	0,9600	48,000
I	50,000	20,000	960	25,000	48,000	64	375,000	1,0000	48,000
J	59,940	16,683	768	29,970	46,034	64	359,640	0,9590	48,000
K	60,000	16,667	768	30,000	46,080	64	360,000	0,9600	48,000
ℓ	120,000	8,333	384	60,000	46,080	64	360,000	0,9600	48,000
M	25,000	40,000	3840	12,500	96,000	128	375,000	1,0000	96,000

Три выделенных столбца в таблице 1 содержат значения управляемых величин, тогда как оставшиеся величины можно рассматривать как зависящие от них. Кроме того, отметим, что идеальными значениями коэффициента передискретизации (SRC) являются $(24/25) \times (1000/1001) \approx 0,9560$, $24/25 = 0,96$ и $1000/1001 \approx 0,9990$. Перечисленные в таблице 1 значения коэффициента SRC округляются, так как являются значениями частоты кадров. Коэффициент 1,000 передискретизации является точным и соответствует отключенному или полностью отсутствующему SRC 130. В примерных вариантах осуществления система 100 обработки аудио работает по меньшей мере в двух режимах с разными длинами кадров, одна или несколько из которых могут совпадать с записями в таблице 1.

Режимы а-d, в которых длина кадра у компонента предварительной обработки устанавливается в 1920 выборок, используются для обращения с частотами (аудио) кадров 23,976, 24,000, 24,975 и 25,000 Гц, выбранными точно совпадающими с частотами видеок кадров широко распространенных форматов кодирования. Из-за разных длин кадров внутренняя частота дискретизации (частота кадров $\times$ длина кадра) будет меняться примерно от 46,034 кГц до 48,000 кГц в режимах а-d; предполагая требовательную дискретизацию и равноотстоящие элементы разрешения по частоте, это будет соответствовать значениям ширины элемента разрешения в диапазоне от 11,988 Гц до 12,500 Гц (половина внутренней частоты дискретизации/длины кадра). Так как колебание внутренних частот дискретизации ограничивается (составляет около 5% как следствие

диапазона колебания частот кадров около 5%), считается, что система 100 добьется приемлемого выходного качества во всех четырех режимах a-d, несмотря на неточное совпадение физической частоты дискретизации, для которой был подготовлен входящий аудиопоток битов.

5 Двигаясь дальше компонента 110 предварительной обработки, гребенка 122 фильтров анализа (QMF) во всех режимах a-d имеет 64 полосы, или 30 выборок на кадр QMF. В физическом выражении это будет соответствовать незначительно меняющейся ширине каждой полосы частот анализа, но колебание опять настолько ограничено, что им можно пренебречь; в частности, модули 124, 126 обработки SBR и DRC могут быть
10 безразличны к текущему режиму без ущерба выходному качеству. Однако SRC 130 зависит от режима и будет использовать специальный коэффициент передискретизации - выбранный для совпадения с частным целевой внешней частоты дискретизации и внутренней частоты дискретизации - чтобы гарантировать, что каждый кадр обработанного аудиосигнала будет содержать некоторое количество выборок,
15 соответствующих целевой внешней частоте дискретизации 48 кГц в физических единицах.

В каждом из режимов a-d система 100 обработки аудио будет точно соответствовать частоте видеок кадров и внешней частоте дискретизации. Система 100 обработки аудио тогда может обрабатывать аудиочасти мультимедийных потоков T1 и T2 битов, где аудиок кадры A11, A12, A13, ...; A22, A23, A24 ... и видеок кадры V11, V12, V13, ...; V22,
20 V23, V24 совпадают по времени в каждом потоке. Тогда возможно улучшить синхронность потоков T1, T2 путем удаления аудиок кадра и ассоциированного видеок кадра в опережающем потоке. В качестве альтернативы аудиок кадр и ассоциированный видеок кадр в отстающем потоке дублируются и вставляются после исходного положения, по возможности совместно с интерполяционными действиями,
25 чтобы уменьшить заметные артефакты.

Режимы e и f, предназначенные для обращения с частотами кадров 29,97 Гц и 30,00 Гц, можно рассматривать как вторую подгруппу. Как уже объяснялось, квантование аудиоданных приспособлено (или оптимизировано) для внутренней частоты дискретизации около 48 кГц. Соответственно, поскольку каждый кадр короче, длина
30 кадра в компоненте 110 предварительной обработки устанавливается в меньшее значение (1536 выборок), чтобы получились внутренние частоты дискретизации около 46,034 и 46,080 кГц. Если гребенка 122 фильтров анализа не зависит от режима и имеет 64 полос частот, то каждый кадр QMF будет содержать 24 выборки.

Аналогичным образом частоты кадров в 50 Гц и 60 Гц или около того
35 (соответствующие удвоенной частоте обновления в стандартизованных телевизионных форматах) и 120 Гц охвачены соответственно режимами g-i (длина кадра 960 выборок), режимами j-k (длина кадра 768 выборок) и режимом l (длина кадра 384 выборок). Отметим, что внутренняя частота дискретизации в каждом случае остается около 48 кГц, чтобы любая психоакустическая подстройка процесса квантования, с помощью
40 которого создавался аудиопоток битов, оставалась по меньшей мере приблизительно верной. Соответствующими длинами кадров QMF в 64-полосной гребенке фильтров будут 15, 12 и 6 выборок.

Как упоминалось, система 100 обработки аудио может работать для подразделения аудиок кадров на более короткие субкадры; причиной для этого может быть более
45 эффективный захват аудиопереходов. Для частоты дискретизации 48 кГц и настроек, приведенных в таблице 1, таблицы 2-4 ниже показывают ширины элементов разрешения и длины кадров, получающиеся в результате подразделения на 2, 4, 8 и 16 субкадров. Считают, что настройки в соответствии с таблицей 1 добиваются выгодного равновесия

разрешения по времени и по частоте.

Таблица 2 Разрешение по времени/частоте при длине кадра в 2048 выборки					
	Количество субкадров				
	1	2	4	8	16
Количество элементов разрешения	2048	1024	512	256	128
Ширина элемента разрешения [Гц]	11,72	23,44	46,88	93,75	187,50
Длительность кадра [мс]	42,67	21,33	10,67	5,33	2,67

Таблица 3 Разрешение по времени/частоте при длине кадра в 1920 выборок					
	Количество субкадров				
	1	2	4	8	16
Количество элементов разрешения	1920	960	480	240	120
Ширина элемента разрешения [Гц]	12,50	25,00	50,00	100,00	200,00
Длительность кадра [мс]	40,00	20,00	10,00	5,00	2,50

Таблица 4 Разрешение по времени/частоте при длине кадра в 1536 выборок					
	Количество субкадров				
	1	2	4	8	16
Количество элементов разрешения	1536	768	384	192	96
Ширина элемента разрешения [Гц]	15,63	31,25	62,50	125,00	250,00
Длительность кадра [мс]	32,00	16,00	8,00	4,00	2,00

Решения, относящиеся к подразделению кадра, могут приниматься как часть процесса подготовки аудиопотока битов, например в системе аудиокодирования (не показана).

Как проиллюстрировано режимом m в таблице 1, систему 100 обработки аудио дополнительно можно приспособить для работы с увеличенной внешней частотой дискретизации 96 кГц и с 128 полосами QMF, соответствующими 30 выборкам на кадр QMF. Поскольку внешняя частота дискретизации, между прочим, совпадает с внутренней частотой дискретизации, коэффициент SRC равен единице, что соответствует отсутствию необходимости передискретизации.

Многоканальное кодирование

При использовании в этом разделе аудиосигнал может быть чистым аудиосигналом, аудиочастью аудиовизуального сигнала либо мультимедийного сигнала или любым из них совместно с метаданными.

При использовании в этом разделе понижающее микширование множества сигналов означает объединение множества сигналов, например, путем образования линейных комбинаций, так что получается меньшее количество сигналов. Обратная к понижающему микшированию операция называется повышающим микшированием, то есть выполнением операции над меньшим количеством сигналов для получения большего количества сигналов.

Фиг.7 - обобщенная блок-схема декодера 100 в системе обработки многоканального аудио для восстановления M кодированных каналов. Декодер 100 содержит три концептуальные части 200, 300, 400, которые будут подробнее объясняться ниже в сочетании с фиг.17-19. В первой концептуальной части 200 кодер принимает N кодированных по форме сигналов понижающего микширования и M кодированных по форме сигналов, представляющих многоканальный аудиосигнал для декодирования,

где $1 < N < M$. В проиллюстрированном примере N устанавливается в 2. Во второй концептуальной части 300 M кодированных по форме сигналов подвергаются понижающему микшированию и объединяются с N кодированными по форме сигналами понижающего микширования. Затем выполняется высокочастотное восстановление (HFR) для объединенных сигналов понижающего микширования. В третьей концептуальной части 400 высокочастотно восстановленные сигналы подвергаются повышающему микшированию, и M кодированных по форме сигналов объединяются с сигналами повышающего микширования, чтобы восстановить M кодированных каналов.

В примерном варианте осуществления, описанном в сочетании с фиг.8-10, описывается восстановление кодированного окружающего звука формата 5.1. Можно отметить, что сигнал низкочастотных эффектов не упоминается в описанном варианте осуществления или на чертежах. Это не означает, что пренебрегают любыми низкочастотными эффектами. Низкочастотные эффекты (Lfe) добавляются в восстановленные 5 каналов любым подходящим способом, известным специалисту в данной области техники. Также можно отметить, что описанный декодер в равной степени подходит для других типов кодированного окружающего звука, например окружающего звука формата 7.1 или 9.1.

Фиг.8 иллюстрирует первую концептуальную часть 200 декодера 100 на фиг.7. Декодер содержит два приемных каскада 212, 214. В первом приемном каскаде 212 поток 202 битов декодируется и деквантуется на два кодированных по форме сигнала 208a-b понижающего микширования. Каждый из двух кодированных по форме сигналов 208a-b понижающего микширования содержит спектральные коэффициенты, соответствующие частотам между первой переходной частотой k_y и второй переходной частотой k_x .

Во втором приемном каскаде 214 поток 202 битов декодируется и деквантуется на пять кодированных по форме сигналов 210a-e. Каждый из пяти кодированных по форме сигналов 210a-e понижающего микширования содержит спектральные коэффициенты, соответствующие частотам вплоть до первой переходной частоты k_x .

В качестве примера сигналы 210a-e содержат два элемента канальной пары и всего один элемент канала для центрального канала. Элементы канальной пары могут быть, например, сочетанием левого переднего и левого окружающего сигнала и сочетанием правого переднего и правого окружающего сигнала. Дополнительным примером является сочетание левого переднего и правого переднего сигналов и сочетание левого окружающего и правого окружающего сигнала. Эти элементы канальной пары могут кодироваться, например, в суммарно-разностном формате. Все пять сигналов 210a-e можно кодировать с использованием перекрывающихся кадрированных преобразований с независимым кадрированием, и их все же можно декодировать с помощью декодера. Это может предусматривать повышенное качество кодирования и, соответственно, повышенное качество декодированного сигнала.

В качестве примера первая переходная частота k_y равна 1,1 кГц. В качестве примера вторая переходная частота k_x находится в диапазоне 5,6-8 кГц. Следует отметить, что первая переходная частота k_y может меняться, даже на основе отдельного сигнала, то есть кодер может обнаружить, что составляющую сигналу в определенном выходном сигнале нельзя точно воспроизвести стереофоническими сигналами 208a-b понижающего микширования, и для того конкретного момента времени может увеличить полосу пропускания, то есть первую переходную частоту k_y , у соответствующего кодированного

по форме сигнала, то есть 210а-е, для надлежащего кодирования той составляющей сигнала по форме.

Как будет позже описываться в этом описании, оставшиеся каскады кодера 100 обычно работают в области квадратурных зеркальных фильтров (QMF). По этой причине каждый из сигналов 208а-б, 210а-е, принятых первым и вторым приемным каскадом 212, 214, которые принимаются в виде модифицированного дискретного косинусного преобразования (MDCT), преобразуется во временную область путем применения обратного MDCT 216. Затем каждый сигнал обратно преобразуется в частотную область путем применения преобразования 218 QMF.

На фиг.9 пять кодированных по форме сигналов 210 подвергаются понижающему микшированию до двух сигналов 310, 312 понижающего микширования, содержащих спектральные коэффициенты, соответствующие частотам вплоть до первой переходной частоты k_y , в каскаде 308 понижающего микширования. Эти сигналы 310, 312 понижающего микширования можно образовать путем выполнения понижающего микширования над низкочастотными многоканальными сигналами 210а-е, используя такую же схему понижающего микширования, которая использовалась в кодере, чтобы создать два сигнала 208а-б понижающего микширования, показанные на фиг.8.

Два новых сигнала 310, 312 понижающего микширования затем объединяются в первом объединяющем каскаде 320, 322 с соответствующим сигналом 208а-б понижающего микширования, чтобы образовать объединенные сигналы 302а-б понижающего микширования. Каждый из объединенных сигналов 302а-б понижающего микширования соответственно содержит спектральные коэффициенты, соответствующие частотам вплоть до первой переходной частоты k_y , возникающим из сигналов 310, 312 понижающего микширования, и спектральные коэффициенты, соответствующие частотам между первой переходной частотой k_y и второй переходной частотой k_x , возникающим из двух кодированных по форме сигналов 208а-б понижающего микширования, принятых в первом приемном каскаде 212 (показаны на фиг.8).

Кодер дополнительно содержит каскад 314 высокочастотного восстановления (HFR). Каскад HFR конфигурируется для расширения каждого из двух объединенных сигналов 302а-б понижающего микширования из объединяющего каскада до частотного диапазона выше второй переходной частоты k_x путем выполнения высокочастотного восстановления. Выполняемое высокочастотное восстановление в соответствии с некоторыми вариантами осуществления может содержать выполнение копирования спектральных полос, SBR. Высокочастотное восстановление может выполняться с использованием параметров высокочастотного восстановления, которые могут быть приняты каскадом 314 HFR любым подходящим способом.

Результатом каскада 314 высокочастотного восстановления являются два сигнала 304а-б, содержащие сигналы 208а-б понижающего микширования с примененным расширением 316, 318 HFR. Как описано выше, каскад 314 HFR выполняет высокочастотное восстановление на основе частот, присутствующих во входном сигнале 210а-е из второго приемного каскада 214 (показан на фиг.8), объединенном с двумя сигналами 208а-б понижающего микширования. Отчасти упрощая, диапазон 316, 318 HFR содержит части спектральных коэффициентов из сигналов 310, 312 понижающего микширования, которые скопированы в диапазон 316, 318 HFR. Следовательно, части пяти кодированных по форме сигналов 210а-е появятся в диапазоне 316, 318 HFR результата 304 из каскада 314 HFR.

Следует отметить, что понижающее микширование в каскаде 308 понижающего

микширования и объединение в первом объединяющем каскаде 320, 322 перед каскадом 314 высокочастотного восстановления могут выполняться во временной области, то есть после того, как каждый сигнал преобразован во временную область путем применения обратного модифицированного дискретного косинусного преобразования 216 (MDCT) (показано на фиг.8). Однако с учетом того, что кодированные по форме сигналы 210a-e и кодированные по форме сигналы 208a-b понижающего микширования могут кодироваться кодировщиком по форме сигнала, использующим перекрывающиеся кадрированные преобразования с независимым кадрированием, сигналы 210a-e и 208a-b нельзя плавно объединить во временной области. Таким образом, достигается сценарий с лучшим управлением, если по меньшей мере объединение в первом объединяющем каскаде 320, 322 выполняется в области QMF.

Фиг.10 иллюстрирует третью и завершающую концептуальную часть 400 декодера 100. Результат 304 каскада 314 HFR образует входные данные в каскад 402 повышающего микширования. Каскад 402 повышающего микширования создает пять выходов 404a-e сигнала путем выполнения параметрического повышающего микширования над расширенными по частоте сигналами 304a-b. Каждый из пяти сигналов 404a-e повышающего микширования соответствует одному из пяти кодированных каналов в кодированном окружающем звуке формата 5.1 для частот выше первой переходной частоты k_y . В соответствии с примерной процедурой параметрического повышающего микширования каскад 402 повышающего микширования сначала принимает параметры параметрического микширования. Каскад 402 повышающего микширования дополнительно формирует декоррелированные версии двух расширенных по частоте, объединенных сигналов 304a-b понижающего микширования. Каскад 402 повышающего микширования дополнительно подвергает матричной операции два расширенных по частоте, объединенных сигнала 304a-b понижающего микширования и декоррелированные версии двух расширенных по частоте, объединенных сигналов 304a-b понижающего микширования, где параметры матричной операции задаются параметрами повышающего микширования. В качестве альтернативы может применяться любая другая процедура параметрического повышающего микширования, известная в данной области техники. Применимые процедуры параметрического повышающего микширования описываются, например, в "MPEG Surround-The ISO/MPEG Standard for Efficient and Compatible Multichannel Audio Coding" (Herre и др., Journal of the Audio Engineering Society, том 56, № 11, ноябрь 2008 г.).

Результат 404a-e каскада 402 повышающего микширования, таким образом, не содержит частоты ниже первой переходной частоты k_y . Оставшиеся спектральные коэффициенты, соответствующие частотам вплоть до первой переходной частоты k_y , находятся в пяти кодированных по форме сигналах 210a-e, которые задержаны каскадом 412 задержки для синхронизации с сигналами 404 повышающего микширования.

Кодер 100 дополнительно содержит второй объединяющий каскад 416, 418. Второй объединяющий каскад 416, 418 конфигурируется для объединения пяти сигналов 404a-e повышающего микширования с пятью кодированными по форме сигналами 210a-e, которые принимались вторым приемным каскадом 214 (показано на фиг.8).

Можно отметить, что в результирующий объединенный сигнал 422 может добавляться любой присутствующий сигнал Lfe как отдельный сигнал. Каждый из сигналов 422 затем преобразуется во временную область путем применения обратного преобразования 420 QMF. Таким образом, результатом обратного преобразования 414

QMF является полностью декодированный 5.1-канальный аудиосигнал.

Фиг.11 иллюстрирует систему 100' декодирования, являющуюся модификацией системы 100 декодирования из фиг.7. Система 100' декодирования имеет концептуальные части 200', 300' и 400', соответствующие концептуальным частям 100, 200 и 300 из фиг.16. Отличие между системой 100' декодирования из фиг.11 и системой декодирования из фиг.7 состоит в том, что имеется третий приемный каскад 616 в концептуальной части 200' и перемежающий каскад 714 в третьей концептуальной части 400'.

Третий приемный каскад 616 конфигурируется для приема дополнительного кодированного по форме сигнала. Дополнительный кодированный по форме сигнал содержит спектральные коэффициенты, соответствующие подмножеству частот выше первой переходной частоты. Дополнительный кодированный по форме сигнал можно преобразовать во временную область путем применения обратного MDCT 216. Затем его можно обратно преобразовать в частотную область путем применения преобразования 218 QMF.

Нужно понимать, что дополнительный кодированный по форме сигнал можно принимать как отдельный сигнал. Однако дополнительный кодированный по форме сигнал также может образовывать часть одного или нескольких из пяти кодированных по форме сигналов 210a-e. Другими словами, дополнительный кодированный по форме сигнал может совместно кодироваться с одним или несколькими из пяти кодированных по форме сигналов 201a-e, например, используя такое же преобразование MCDT. Если это имеет место, то третий приемный каскад 616 соответствует второму приемному каскаду, то есть дополнительный кодированный по форме сигнал принимается вместе с пятью кодированными по форме сигналами 210a-e через второй приемный каскад 214.

Фиг.12 подробнее иллюстрирует третью концептуальную часть 300' декодера 100' из фиг.11. Дополнительный кодированный по форме сигнал 710 вводится в третью концептуальную часть 400' в дополнение к высокочастотным расширенным сигналам 304a-b понижающего микширования и пяти кодированным по форме сигналам 210a-e. В проиллюстрированном примере дополнительный кодированный по форме сигнал 710 соответствует третьему каналу из пяти каналов. Дополнительный кодированный по форме сигнал 710 дополнительно содержит спектральные коэффициенты, соответствующие частотному интервалу, который начинается с первой переходной частоты k_y . Однако вид подмножества частотного диапазона выше первой переходной частоты, охваченного дополнительным кодированным по форме сигналом 710, конечно, может меняться в разных вариантах осуществления. Также нужно отметить, что может быть принято множество кодированных по форме сигналов 710a-e, где разные кодированные по форме сигналы могут соответствовать разным выходным каналам. Подмножество частотного диапазона, охваченное множеством дополнительных кодированных по форме сигналов 710a-e, может меняться между разными сигналами в множестве дополнительных кодированных по форме сигналов 710a-e.

Дополнительный кодированный по форме сигнал 710 может быть задержан каскадом 712 задержки для синхронизации сигналов 404 повышающего микширования, выводимых из каскада 402 повышающего микширования. Затем сигналы 404 повышающего микширования и дополнительный кодированный по форме сигнал 710 вводятся в перемежающий каскад 714. Перемежающий каскад 714 перемежает, то есть объединяет сигналы 404 повышающего микширования с дополнительным кодированным по форме сигналом 710, чтобы сформировать перемеженный сигнал 704. В настоящем примере перемежающий каскад 714 перемежает, таким образом, третий сигнал 404c

повышающего микширования с дополнительным кодированным по форме сигналом 710. Перемежение может выполняться путем сложения двух сигналов. Однако обычно перемежение выполняется путем замены сигналов 404 повышающего микширования дополнительным кодированным по форме сигналом 710 в частотном диапазоне и временном диапазоне, где сигналы перекрываются.

Перемеженный сигнал 704 затем вводится во второй объединяющий каскад 416, 418, где он объединяется с кодированными по форме сигналами 201a-e, чтобы сформировать выходной сигнал 722 таким же образом, как описано со ссылкой на фиг.19. Нужно отметить, что порядок перемежающего каскада 714 и второго объединяющего каскада 416, 418 можно изменить на противоположный, чтобы объединение выполнялось до перемежения.

Также в ситуации, где дополнительный кодированный по форме сигнал 710 образует часть одного или нескольких из пяти кодированных по форме сигналов 210a-e, второй объединяющий каскад 416, 418 и перемежающий каскад 714 можно объединить в один каскад. В частности, такой объединенный каскад использовал бы спектральное содержимое пяти кодированных по форме сигналов 210a-e для частот вплоть до первой переходной частоты k_y . Для частот выше первой переходной частоты объединенный каскад использовал бы сигналы 404 повышающего микширования, перемеженные с дополнительным кодированным по форме сигналом 710.

Перемежающий каскад 714 может работать под управлением управляющего сигнала. С этой целью декодер 100' может принимать, например, через третий приемный каскад 616, управляющий сигнал, который указывает, как перемежать дополнительный кодированный по форме сигнал с одним из M сигналов повышающего микширования. Например, управляющий сигнал может указывать частотный диапазон и временной диапазон, для которого нужно перемежать дополнительный кодированный по форме сигнал 710 с одним из сигналов 404 повышающего микширования. Например, частотный диапазон и временной диапазон могут выражаться в показателях фрагментов времени/частоты, для которых нужно выполнить перемежение. Фрагменты времени/частоты могут быть фрагментами времени/частоты по отношению к временной/частотной сетке области QMF, где происходит перемежение.

Управляющий сигнал может использовать векторы, например двоичные векторы, для указания фрагментов времени/частоты, для которых нужно выполнить перемежение. В частности, первый вектор, относящийся к направлению частоты, может указывать частоты, для которых нужно выполнить перемежение. Указание может осуществляться, например, путем указания логической единицы для соответствующего частотного интервала в первом векторе. Второй вектор, относящийся к направлению времени, также может указывать интервалы времени, для которых нужно выполнить перемежение. Указание может осуществляться, например, путем указания логической единицы для соответствующего интервала времени во втором векторе. С этой целью временной кадр обычно разделяется на множество временных интервалов, так что указание времени может осуществляться на основе субкадров. С помощью пересечения первого и второго векторов можно построить матрицу времени/частоты. Например, матрица времени/частоты может быть двоичной матрицей, содержащей логическую единицу для каждого фрагмента времени/частоты, для которого первый и второй векторы указывают логическую единицу. Перемежающий каскад 714 тогда может использовать матрицу времени/частоты при выполнении перемежения, например, так, что один или несколько сигналов 704 повышающего микширования заменяются дополнительным кодированным по форме сигналом 710 для фрагментов времени/частоты, указываемых

в матрице времени/частоты, например, с помощью логической единицы.

Отметим, что векторы могут использовать другие схемы, помимо двоичной, для указания фрагментов времени/частоты, для которых нужно выполнить перемежение. Например, векторы могли бы указывать посредством первого значения, например
 5 нуля, что перемежение не нужно выполнять, и второго значения, что перемежение нужно выполнить по отношению к некоторому каналу, идентифицированному вторым значением.

Стереофоническое кодирование

При использовании в этом разделе лево-правое кодирование означает, что левый
 10 (L) и правый (R) стереофонические сигналы кодируются без выполнения какого-либо преобразования между сигналами.

При использовании в этом разделе суммарно-разностное кодирование означает, что сумма M левого и правого стереофонических сигналов кодируется как один сигнал (сумма), и разность S между левым и правым стереофоническими сигналами кодируется
 15 как один сигнал (разность). Суммарно-разностное кодирование также может называться средне-боковым кодированием. Связь между лево-правым видом и суммарно-разностным видом, таким образом, выглядит как $M = L + R$ и $S = L - R$. Можно отметить, что возможны разные нормализации или масштабирование при преобразовании левого и правого стереофонических сигналов в суммарно-разностный вид и наоборот при
 20 условии, что преобразование в обоих направлениях совпадает. В данном раскрытии изобретения в основном используется $M = L + R$ и $S = L - R$, но система, использующая другое масштабирование, например $M = (L + R)/2$ и $S = (L - R)/2$, работает с тем же успехом.

При использовании в этом разделе кодирование с понижающе-дополняющим
 25 микшированием (dmx/comp) означает подвергание левого и правого стереофонического сигнала матричному умножению в зависимости от весового параметра a перед кодированием. Соответственно, кодирование dmx/comp также может называться кодированием dmx/comp/a. Связь между видом с понижающе-дополняющим микшированием, лево-правым видом и суммарно-разностным вид обычно выглядит
 30 как $dmx = L + R = M$, и $comp = (1 - a)L - (1 + a)R = -aM + S$. Примечательно, что сигнал понижающего микширования в представлении с понижающе-дополняющим микшированием эквивалентен соответственно суммарному сигналу M в суммарно-разностном представлении.

При использовании в этом разделе аудиосигнал может быть чистым аудиосигналом,
 35 аудиочастью аудиовизуального сигнала либо мультимедийного сигнала или любым из них совместно с метаданными.

Фиг.13 - обобщенная блок-схема системы 100 декодирования, содержащей три концептуальные части 200, 300, 400, которые будут подробнее объясняться в сочетании с фиг.14-16 ниже. В первой концептуальной части 200 поток битов принимается и
 40 декодируется на первый и второй сигналы. Первый сигнал содержит первый кодированный по форме сигнал, содержащий спектральные данные, соответствующие частотам вплоть до первой переходной частоты, и кодированный по форме сигнал понижающего микширования, содержащий спектральные данные, соответствующие частотам выше первой переходной частоты. Второй сигнал содержит только второй
 45 кодированный по форме сигнал, содержащий спектральные данные, соответствующие частотам вплоть до первой переходной частоты.

Во второй концептуальной части 300, если кодированные по форме части первого и второго сигналов не имеют суммарно-разностный вид, например, имеют вид M/S, то

кодированные по форме части первого и второго сигналов преобразуются к суммарно-разностному виду. После этого первый и второй сигналы преобразуются во временную область, а затем в область квадратурных зеркальных фильтров, QMF. В третьей концептуальной части 400 первый сигнал подвергается высокочастотному
 5 восстановлению (HFR). Первый и второй сигналы затем подвергаются повышающему микшированию, чтобы создать левый и правый выходы стереофонического сигнала, имеющие спектральные коэффициенты, соответствующие всей полосе частот кодированного сигнала, декодируемого системой 100 декодирования.

Фиг.14 иллюстрирует первую концептуальную часть 200 системы 100 декодирования
 10 на фиг.13. Система 100 декодирования содержит приемный каскад 212. В приемном каскаде 212 кадр 202 потока битов декодируется и деквантуется на первый сигнал 204a и второй сигнал 204b. Кадр 202 потока битов соответствует временному кадру двух декодируемых аудиосигналов. Первый сигнал 204a содержит первый кодированный по форме сигнал 208, содержащий спектральные данные, соответствующие частотам
 15 вплоть до первой переходной частоты k_y , и кодированный по форме сигнал 206 понижающего микширования, содержащий спектральные данные, соответствующие частотам выше первой переходной частоты k_y . В качестве примера первая переходная частота k_y равна 1,1 кГц.

Кодированный по форме сигнал 206 понижающего микширования в соответствии
 20 с некоторыми вариантами осуществления содержит спектральные данные, соответствующие частотам между первой переходной частотой k_y и второй переходной частотой k_x . В качестве примера вторая переходная частота k_x находится в диапазоне 5,6-8 кГц.

Принятые первый и второй кодированные по форме сигналы 208, 210 можно
 25 кодировать по форме в лево-правом виде, суммарно-разностном виде и/или виде с понижающе-дополняющим микшированием, где дополняющий сигнал зависит от адаптивного к сигналу весового параметра a . Кодированный по форме сигнал 206 понижающего микширования соответствует понижающему микшированию,
 30 подходящему для параметрического стерео, которое в соответствии с вышеизложенным соответствует суммарному виду. Однако сигнал 204b не имеет содержимого выше первой переходной частоты k_y . Каждый из сигналов 206, 208, 210 представляется в области модифицированного дискретного косинусного преобразования (MDCT).

Фиг.15 иллюстрирует вторую концептуальную часть 300 системы 100 декодирования
 35 на фиг.13. Система 100 декодирования содержит каскад 302 микширования. Исполнение системы 100 декодирования требует, чтобы вход в каскад высокочастотного восстановления, который будет подробнее описываться ниже, был в формате суммы. Следовательно, каскад микширования конфигурируется для проверки, имеют ли
 40 суммарно-разностный вид первый и второй кодированные по форме сигналы 208, 210. Если первый и второй кодированные по форме сигналы 208, 210 не имеют суммарно-разностный вид для всех частот вплоть до первой переходной частоты k_y , то каскад 302 микширования преобразует все кодированные по форме сигналы 208, 210 в суммарно-разностный вид. Если по меньшей мере подмножество частот у входных
 45 сигналов 208, 210 в каскад 302 микширования имеет вид с понижающе-дополняющим микшированием, то необходим весовой параметр a в качестве входа в каскад 302 микширования. Можно отметить, что входные сигналы 208, 210 могут содержать несколько подмножеств частот, кодированных в виде с понижающе-дополняющим микшированием, и что в этом случае не нужно кодировать каждое подмножество с

использованием одного и того же значения весового параметра a . В этом случае необходимо несколько весовых параметров a в качестве входа в каскад 302 микширования.

Как упоминалось выше, каскад 302 микширования всегда выводит суммарно-разностное представление входных сигналов 204a-b. Чтобы иметь возможность преобразовать в суммарно-разностное представление сигналы, представленные в области MDCT, кадрирование кодированных по MDCT сигналов должно быть одинаковым. Это подразумевает, что если первый и второй кодированные по форме сигналы 208, 210 имеют вид L/R или вид с понижающе-дополняющим микшированием, то кадрирование для сигнала 204a и кадрирование для сигнала 204b не может быть независимым.

Следовательно, если первый и второй кодированные по форме сигналы 208, 210 имеют суммарно-разностный вид, то кадрирование для сигнала 204a и кадрирование для сигнала 204b может быть независимым.

После каскада 302 микширования суммарно-разностный сигнал преобразуется во временную область путем применения обратного модифицированного дискретного косинусного преобразования 312 (MDCT⁻¹).

Затем два сигнала 304a-b анализируются с помощью двух гребенок 314 QMF. Поскольку сигнал 306 понижающего микширования не содержит низкие частоты, не нужно анализировать сигнал с помощью гребенки фильтров Найквиста, чтобы повысить разрешение по частоте. Это можно сравнить с системами, где сигнал понижающего микширования содержит низкие частоты, например традиционное параметрическое стереофоническое декодирование, такое как параметрическое стерео MPEG-4. В тех системах сигнал понижающего микширования нужно анализировать с помощью гребенки фильтров Найквиста, чтобы повысить разрешение по частоте сверх того, что достигается гребенкой QMF, и соответственно лучше соответствует избирательности по частоте у слуховой системы человека, которая представлена, например, шкалой барков.

Выходной сигнал 304 из гребенок 314 QMF содержит первый сигнал 304a, который является сочетанием кодированного по форме суммарного сигнала 308, содержащего спектральные данные, соответствующие частотам вплоть до первой переходной частоты k_y , и кодированного по форме сигнала 306 понижающего микширования, содержащего спектральные данные, соответствующие частотам между первой переходной частотой k_y и второй переходной частотой k_x . Выходной сигнал 304 дополнительно содержит второй сигнал 304b, который содержит кодированный по форме разностный сигнал 310, содержащий спектральные данные, соответствующие частотам вплоть до первой переходной частоты k_y . Сигнал 304b не имеет содержимого выше первой переходной частоты k_y .

Как будет описываться позже, каскад 416 высокочастотного восстановления (показан в сочетании с фиг.16) использует низкие частоты, то есть первый кодированный по форме сигнал 308 и кодированный по форме сигнал 306 понижающего микширования из выходного сигнала 304, для восстановления частот выше второй переходной частоты k_x . Выгодно, чтобы сигнал, на который воздействует каскад 416 высокочастотного восстановления, был сигналом аналогичного типа на низких частотах. С этой точки зрения выгодно заставить каскад 302 микширования всегда выводить суммарно-разностное представление первого и второго кодированных по форме сигналов 208, 210, поскольку это подразумевает, что первый кодированный по форме сигнал 308 и

кодированный по форме сигнал 306 понижающего микширования из выведенного первого сигнала 304а обладают сходным характером.

Фиг.16 иллюстрирует третью концептуальную часть 400 системы 100 декодирования на фиг.13. Каскад 416 высокочастотного восстановления (HFR) расширяет сигнал 306 понижающего микширования в первом входном сигнале 304а до частотного диапазона выше второй переходной частоты k_x путем выполнения высокочастотного восстановления. В зависимости от конфигурации каскада 416 HFR входом в каскад 416 HFR является весь сигнал 304а или только сигнал 306 понижающего микширования. Высокочастотное восстановление выполняется с использованием параметров высокочастотного восстановления, которые могут быть приняты каскадом 416 высокочастотного восстановления любым подходящим способом. Выполняемое высокочастотное восстановление в соответствии с вариантом осуществления содержит выполнение копирования спектральных полос, SBR.

Результатом каскада 314 высокочастотного восстановления является сигнал 404, содержащий сигнал 406 понижающего микширования с примененным расширением 412 SBR. Высокочастотно восстановленный сигнал 404 и сигнал 304b затем подаются в каскад 420 повышающего микширования, чтобы сформировать левый L и правый R стереофонические сигналы 412а-b. Для спектральных коэффициентов, соответствующих частотам ниже первой переходной частоты k_y , повышающее микширование содержит выполнение обратного суммарно-разностного преобразования первого и второго сигналов 408, 310. Это просто означает переход из средне-бокового представления в лево-правое представление, как указывалось раньше. Для спектральных коэффициентов, соответствующих частотам сверх первой переходной частоты k_y , сигнал 406 понижающего микширования и расширение 412 SBR подаются через декоррелятор 418. Сигнал 406 понижающего микширования и расширение 412 SBR, и декоррелированная версия сигнала 406 понижающего микширования и расширения 412 SBR затем подвергаются повышающему микшированию с использованием параметров параметрического микширования, чтобы восстановить левый и правый каналы 416, 414 для частот выше первой переходной частоты k_y . Может применяться любая процедура параметрического повышающего микширования, известная в данной области техники.

Следует отметить, что в вышеприведенном примерном варианте 100 осуществления кодера, показанном на фиг.13-16, необходимо высокочастотное восстановление, поскольку первый принятый сигнал 204а содержит только спектральные данные, соответствующие частотам вплоть до второй переходной частоты k_x . В дополнительных вариантах осуществления первый принятый сигнал содержит спектральные данные, соответствующие всем частотам кодированного сигнала. В соответствии с этим вариантом осуществления высокочастотное восстановление не нужно. Специалист в данной области техники понимает, как в этом случае приспособить примерный кодер 100.

Фиг.17 в качестве примера показывает обобщенную блок-схему системы 500 кодирования в соответствии с вариантом осуществления.

В системе кодирования первый и второй сигналы 540, 542 для кодирования принимаются приемным каскадом (не показан). Эти сигналы 540, 542 представляют временной кадр левого 540 и правого 542 стереофонических аудиоканалов. Сигналы 540, 542 представляются во временной области. Система кодирования содержит каскад 510 преобразования. Сигналы 540, 542 преобразуются в суммарно-разностный формат

544, 546 в каскаде 510 преобразования.

Система кодирования дополнительно содержит каскад 514 кодирования по форме, сконфигурированный для приема первого и второго преобразованных сигналов 544, 546 из каскада 510 преобразования. Каскад кодирования по форме обычно работает в области MDCT. По этой причине преобразованные сигналы 544, 546 подвергаются преобразованию 512 MDCT перед каскадом 514 кодирования по форме. В каскаде кодирования по форме первый и второй преобразованные сигналы 544, 546 кодируются по форме соответственно в первый и второй кодированные по форме сигналы 518, 520.

Для частот выше первой переходной частоты k_y каскад 514 кодирования по форме конфигурируется для кодирования по форме первого преобразованного сигнала 544 в кодированный по форме сигнал 552 в первом кодированном по форме сигнале 518. Каскад 514 кодирования по форме может конфигурироваться для установки второго кодированного по форме сигнала 520 в ноль при превышении первой переходной частоты k_y или вообще для отказа от кодирования этих частот. Для частот выше первой переходной частоты k_y каскад 514 кодирования по форме конфигурируется для кодирования по форме первого преобразованного сигнала 544 в кодированный по форме сигнал 552 в первом кодированном по форме сигнале 518.

Для частот ниже первой переходной частоты k_y в каскаде 514 кодирования по форме принимается решение о том, какой вид стереофонического кодирования использовать для двух сигналов 548, 550. В зависимости от характеристик преобразованных сигналов 544, 546 ниже первой переходной частоты k_y могут приниматься разные решения для разных подмножеств кодированного по форме сигнала 548, 550. Кодирование может быть либо лево-правым кодированием, средне-боковым кодированием, то есть кодированием суммы и разности, либо кодированием dmx/comp/a. Если сигналы 548, 550 кодируются по форме с помощью суммарно-разностного кодирования в каскаде 514 кодирования по форме, то кодированные по форме сигналы 518, 520 могут кодироваться с использованием перекрывающихся кадрированных преобразований с независимым кадрированием для сигналов 518, 520 соответственно.

Примерная первая переходная частота k_y равна 1,1 кГц, но эта частота может меняться в зависимости от скорости передачи битов у стереофонической аудиосистемы или в зависимости от характеристик аудио, которое нужно кодировать.

По меньшей мере два сигнала 518, 520 выводятся соответственно из каскада 514 кодирования по форме. Если одно или несколько подмножеств или вся полос частот сигналов ниже первой переходной частоты k_y кодируются в виде с понижающе-дополняющим микшированием путем выполнения матричной операции в зависимости от весового параметра a , то этот параметр также выводится как сигнал 522. В случае нескольких подмножеств, кодируемых в виде с понижающе-дополняющим микшированием, каждое подмножество не нужно кодировать с использованием одинакового значения весового параметра a . В этом случае несколько весовых параметров выводятся как сигнал 522.

Эти два или три сигнала 518, 520, 522 кодируются и квантуются 524 в одиночный полный сигнал 558.

Чтобы на стороне декодера иметь возможность восстанавливать спектральные данные из первого и второго сигналов 540, 542 для частот выше первой переходной частоты, нужно извлечь параметры 536 параметрического стерео из сигналов 540, 542. С этой целью кодер 500 содержит каскад 530 параметрического стереофонического (PS) кодирования. Каскад 530 PS-кодирования обычно работает в области QMF. Поэтому

перед вводом в каскад 530 PS-кодирования первый и второй сигналы 540, 542 преобразуются в область QMF с помощью каскада 526 анализа QMF. Каскад 530 PS-кодирования приспособлен только для извлечения параметров 536 параметрического стерео для частот выше первой переходной частоты k_y .

Можно отметить, что параметры 536 параметрического стерео отражают характеристики сигнала, который подвергается параметрическому стереофоническому кодированию. Соответственно, они избирательны по частоте, то есть каждый параметр из параметров 536 может соответствовать подмножеству частот левого или правого входного сигнала 540, 542. Каскад 530 PS-кодирования вычисляет параметры 536 параметрического стерео и квантует эти параметры либо равномерно, либо неравномерно. Параметры, как упоминалось выше, вычисляются избирательными по частоте, где весь частотный диапазон входных сигналов 540, 542 разделяется, например, на 15 параметрических полос. Они могут быть разнесены в соответствии с моделью разрешения по частоте у слуховой системы человека, например шкалой барков.

В примерном варианте осуществления кодера 500, показанном на фиг.17, каскад 514 кодирования по форме конфигурируется для кодирования по форме первого преобразованного сигнала 544 для частот между первой переходной частотой k_y и второй переходной частотой k_x и установки первого кодированного по форме сигнала 518 в ноль при превышении второй переходной частоты k_x . Это может выполняться для дальнейшего уменьшения необходимой скорости передачи в аудиосистеме, частью которой является кодер 500. Чтобы иметь возможность восстанавливать сигнал выше второй переходной частоты k_x , нужно формировать параметры 538 высокочастотного восстановления. В соответствии с этим примерным вариантом осуществления это осуществляется путем понижающего микширования двух сигналов 540, 542, представленных в области QMF, в каскаде 534 понижающего микширования. Результирующий сигнал понижающего микширования, который равен, например, сумме сигналов 540, 542, подвергается затем кодированию с высокочастотным восстановлением в каскаде 532 кодирования с высокочастотным восстановлением, HFR, чтобы сформировать параметры 538 высокочастотного восстановления. Параметры 538 могут включать в себя, например, огибающую спектра у частот выше второй переходной частоты k_x , информацию о накоплении помехи т. п., которые известны специалисту в данной области техники.

Примерная вторая переходная частота k_x равна 5,6 - 8 кГц, но эта частота может меняться в зависимости от скорости передачи битов у стереофонической аудиосистемы или в зависимости от характеристик аудио, которое нужно кодировать.

Кодер 500 дополнительно содержит каскад 524 формирования потока битов, то есть мультиплексор потока битов. В соответствии с примерным вариантом осуществления кодера 500 каскад формирования потока битов конфигурируется для приема кодированного и квантованного сигнала 544 и двух сигналов 536, 538 параметров. Они преобразуются в поток 560 битов с помощью каскада 562 формирования потока битов для дальнейшего распространения в стереофонической аудиосистеме.

В соответствии с другим вариантом осуществления каскад 514 кодирования по форме конфигурируется для кодирования по форме первого преобразованного сигнала 544 для всех частот выше первой переходной частоты k_y . В этом случае каскад 532 кодирования с HFR не нужен, и следовательно, в поток битов не включаются никакие параметры 538 высокочастотного восстановления.

Фиг.18 в качестве примера показывает обобщенную блок-схему системы 600 кодирования в соответствии с другим вариантом осуществления.

Кодирование в речевом режиме

Фиг.19а показывает блок-схему примерного речевого кодера 100 с преобразованием.

5 Кодер 100 в качестве входа принимает блок 131 коэффициентов преобразования (также называемый единицей кодирования). Блок 131 коэффициентов преобразования может быть получен узлом преобразования, сконфигурированным для преобразования последовательности выборок входного аудиосигнала из временной области в область преобразования. Узел преобразования может конфигурироваться для выполнения

10 MDCT. Узел преобразования может быть частью универсального аудиокодека, например AAC или HE-AAC. Такой универсальный аудиокодек может применять разные размеры блоков, например длинный блок и короткий блок. Примерные размеры блоков составляют 1024 выборок для длинного блока и 256 выборок для короткого блока. Предполагая частоту дискретизации 44,1 кГц и перекрытие в 50%, длинный блок

15 охватывает приблизительно 20 мс входного аудиосигнала, а короткий блок охватывает приблизительно 5 мс входного аудиосигнала. Длинные блоки обычно используются для стационарных сегментов входного аудиосигнала, а короткие блоки обычно используются для переходных сегментов входного аудиосигнала.

Речевые сигналы можно рассматривать как стационарные во временных сегментах

20 около 20 мс. В частности, огибающая спектра речевого сигнала может считаться стационарной во временных сегментах около 20 мс. Чтобы вывести содержательную статистику в области преобразования для таких сегментов 20 мс, может быть полезно предоставить речевому кодеру 100 с преобразованием короткие блоки 131 коэффициентов преобразования (имеющие длину, например, 5 мс). При этом можно

25 использовать множество коротких блоков 131 для выведения статистики касательно временных сегментов, например, по 20 мс (например, временной сегмент длинного блока). Кроме того, имеется преимущество в обеспечении соразмерного разрешения по времени для речевых сигналов.

Поэтому узел преобразования может конфигурироваться для предоставления

30 коротких блоков 131 коэффициентов преобразования, если текущий сегмент входного аудиосигнала классифицируется как речь. Кодер 100 может содержать узел 101 кадрирования, сконфигурированный для извлечения множества блоков 131 коэффициентов преобразования, называемого набором 132 блоков 131. Набор 132 блоков также может называться кадром. В качестве примера набор 132 блоков 131

35 может содержать четыре коротких блока с 256 коэффициентами преобразования, посредством этого охватывая сегмент входного аудиосигнала приблизительно в 20 мс.

Набор 132 блоков может предоставляться узлу 102 оценки огибающей. Узел 102 оценки огибающей может конфигурироваться для определения огибающей 133 на основе набора 132 блоков. Огибающая 133 может основываться на

40 среднеквадратических (RMS) значениях соответствующих коэффициентов преобразования в множестве блоков 131, содержащихся в наборе 132 блоков. Блок 131 обычно предоставляет множество коэффициентов преобразования (например, 256 коэффициентов преобразования) в соответствующем множестве элементов 301 разрешения по частоте (см. фиг.21а). Множество элементов 301 разрешения по частоте

45 можно сгруппировать в множество полос 302 частот. Множество полос 302 частот может выбираться на основе психоакустических соображений. В качестве примера элементы 301 разрешения по частоте можно сгруппировать в полосы 302 частот в соответствии с логарифмической шкалой или шкалой барков. Огибающая 134, которая

определена на основе текущего набора 132 блоков, может содержать множество значений энергии для множества полос 302 частот соответственно. Конкретное значение энергии для конкретной полосы 302 частот можно определить на основе коэффициентов преобразования в блоках 131 набора 132, которые соответствуют элементам 301 разрешения по частоте, попадающим в конкретную полосу 302 частот. Конкретное значение энергии можно определить на основе RMS-значения этих коэффициентов преобразования. По существу, огибающая 133 для текущего набора 132 блоков (называемая текущей огибающей 133) может указывать среднюю огибающую блоков 131 коэффициентов преобразования, содержащихся в текущем наборе 132 блоков, или может указывать среднюю огибающую блоков 132 коэффициентов преобразования, используемых для определения огибающей 133.

Следует отметить, что текущая огибающая 133 может определяться на основе одного или нескольких дополнительных блоков 131 коэффициентов преобразования рядом с текущим набором 132 блоков. Это иллюстрируется на фиг.20, где текущая огибающая 133 (указанная квантованной текущей огибающей 134) определяется на основе блоков 131 текущего набора 132 блоков и на основе блока 201 из набора блоков, предшествующего текущему набору 132 блоков. В проиллюстрированном примере текущая огибающая 133 определяется на основе пяти блоков 131. Принимая во внимание соседние блоки при определении текущей огибающей 133, можно обеспечить непрерывность огибающих у соседних наборов 132 блоков.

При определении текущей огибающей 133 можно присваивать веса коэффициентам преобразования разных блоков 131. В частности, крайние блоки 201, 202, которые принимаются во внимание для определения текущей огибающей 133, могут иметь меньший вес, чем оставшиеся блоки 131. В качестве примера коэффициенты преобразования у крайних блоков 201, 202 могут иметь вес 0,5, где коэффициенты преобразования у других блоков 131 могут иметь вес 1.

Следует отметить, что аналогично рассмотрению блоков 201 предыдущего набора 132 блоков можно рассматривать один или несколько блоков (так называемые упреждающие блоки) в непосредственно следующем наборе 132 блоков для определения текущей огибающей 133.

Значения энергии текущей огибающей 133 можно представить на логарифмической шкале (например, на шкале в дБ). Текущая огибающая 133 может предоставляться в узел 103 квантования огибающей, который конфигурируется для квантования значений энергии текущей огибающей 133. Узел 103 квантования огибающей может предоставлять заранее установленное разрешение квантователя, например разрешение в 3 дБ. Индексы квантования огибающей 133 могут предоставляться в виде данных 161 огибающей в потоке битов, сформированном кодером 100. Кроме того, квантованная огибающая 134, то есть огибающая, содержащая квантованные значения энергии огибающей 133, может предоставляться в узел 104 интерполяции.

Узел 104 интерполяции конфигурируется для определения огибающей для каждого блока 131 в текущем наборе 132 блоков на основе квантованной текущей огибающей 134 и на основе квантованной предыдущей огибающей 135 (которая определена для набора 132 блоков, непосредственно предшествующего текущему набору 132 блоков). Работа узла 104 интерполяции иллюстрируется на фиг.20, 21a и 21b. Фиг.20 показывает последовательность блоков 131 коэффициентов преобразования. Последовательность блоков 131 группируется в следующий наборов 132 блоков, где каждый набор 132 блоков используется для определения квантованной огибающей, например квантованной текущей огибающей 134 и квантованной предыдущей огибающей 135. Фиг.21a

показывает примеры квантованной предыдущей огибающей 135 и квантованной текущей огибающей 134. Как указано выше, огибающие могут указывать спектральную энергию 303 (например, на шкале в дБ). Соответствующие значения 303 энергии квантованной предыдущей огибающей 135 и квантованной текущей огибающей 134 для одной и той же полосы 302 частот можно интерполировать (например, используя линейную интерполяцию), чтобы определить интерполированную огибающую 136. Другими словами, значения 303 энергии конкретной полосы 302 частот можно интерполировать для предоставления значения 303 энергии интерполированной огибающей 136 в конкретной полосе 302 частот.

Следует отметить, что набор блоков, для которых определяются и применяются интерполированные огибающие 136, может отличаться от текущего набора 132 блоков, на основе которого определяется квантованная текущая огибающая 134. Это иллюстрируется на фиг.20, которая показывает сдвинутый набор 332 блоков, который сдвигается по сравнению с текущим набором 132 блоков и который содержит блоки 3 и 4 из предыдущего набора 132 блоков (указанные номерами 203 и 201 ссылок соответственно) и блоки 1 и 2 из текущего набора 132 блоков (указанные номерами 204 и 205 ссылок соответственно). Фактически, интерполированные огибающие 136, определенные на основе квантованной текущей огибающей 134 и на основе квантованной предыдущей огибающей 135, могут обладать повышенной релевантностью для блоков в сдвинутом наборе 332 блоков по сравнению с релевантностью для блоков в текущем наборе 132 блоков.

Поэтому показанные на фиг.21b интерполированные огибающие 136 можно использовать для выравнивания блоков 131 сдвинутого набора 332 блоков. Это показано с помощью фиг.21b совместно с фиг.20. Видно, что интерполированная огибающая 341 из фиг.21b может применяться к блоку 203 из фиг.20, интерполированная огибающая 342 из фиг.21b может применяться к блоку 201 из фиг.20, интерполированная огибающая 343 из фиг.21b может применяться к блоку 204 из фиг.20, и что интерполированная огибающая 344 из фиг.21b (которая соответствует квантованной текущей огибающей 136 в проиллюстрированном примере) может применяться к блоку 205 из фиг.20. По существу, набор 132 блоков для определения квантованной текущей огибающей 134 может отличаться от сдвинутого набора 332 блоков, для которого определяются интерполированные огибающие 136 и к которому применяются интерполированные огибающие 136 (с целью выравнивания). В частности, квантованная текущая огибающая 134 может определяться с использованием некоторого упреждения относительно блоков 203, 201, 204, 205 в сдвинутом наборе 332 блоков, которые нужно выровнять с использованием квантованной текущей огибающей 134. Это полезно с точки зрения непрерывности.

Интерполяция значений 303 энергии для определения интерполированных огибающих 136 иллюстрируется на фиг.21b. Видно, что путем интерполяции между значением энергии квантованной предыдущей огибающей 135 к соответствующему значению энергии квантованной текущей огибающей 134 можно определить значения энергии интерполированных огибающих 136 для блоков 131 в сдвинутом наборе 332 блоков. В частности, для каждого блока 131 в сдвинутом наборе 332 может определяться интерполированная огибающая 136, посредством этого предоставляя множество интерполированных огибающих 136 для множества блоков 203, 201, 204, 205 в сдвинутом наборе 332 блоков. Интерполированная огибающая 136 блока 131 коэффициентов преобразования (например, любого из блоков 203, 201, 204, 205 в сдвинутом наборе 332 блоков) может использоваться для кодирования блока 131 коэффициентов

преобразования. Следует отметить, что индексы 161 квантования текущей огибающей 133 предоставляются соответствующему декодеру в потоке битов. Следовательно, соответствующий декодер может конфигурироваться для определения множества интерполированных огибающих 136 аналогично узлу 104 интерполяции в кодере 100.

5 Узел 101 кадрирования, узел 103 оценки огибающей, узел 103 квантования огибающей и узел 104 интерполяции воздействуют на набор блоков (то есть текущий набор 132 блоков и/или сдвинутый набор 332 блоков). С другой стороны, фактическое кодирование коэффициента преобразования может выполняться поблочно. Ниже приводится ссылка на кодирование текущего блока 131 коэффициентов преобразования, который может
10 быть любым из множества блоков 131 сдвинутого набора 332 блоков (или, возможно, текущего набора 132 блоков в другой реализации речевого кодера 100 с преобразованием).

Текущая интерполированная огибающая 136 для текущего блока 131 может предоставлять приближение огибающей спектра у коэффициентов преобразования
15 текущего блока 131. Кодер 100 может содержать узел 105 предварительного выравнивания и узел 106 определения усиления огибающей, которые конфигурируются для определения отрегулированной огибающей 139 для текущего блока 131 на основе текущей интерполированной огибающей 136 и на основе текущего блока 131. В частности, усиление огибающей для текущего блока 131 может определяться так, что
20 регулируется дисперсия выровненных коэффициентов преобразования текущего блока 131. $X(k)$, $k = 1, \dots, K$, могут быть коэффициентами преобразования текущего блока 131 (например, при $K = 256$), и $E(k)$, $k = 1, \dots, K$, могут быть средними спектральными значениями 303 энергии у текущей интерполированной огибающей 136 (при равных значениях $E(k)$ энергии у одной и той же полосы 302 частот). Усиление α огибающей
25 может определяться так, что регулируется дисперсия $\tilde{X}(k) = \frac{X(k)}{\alpha \cdot \sqrt{E(k)}}$ выровненных коэффициентов преобразования. В частности, усиление α огибающей может определяться так, что дисперсия равна единице.

Следует отметить, что усиление α огибающей может определяться для субполосы в
30 полном частотном диапазоне текущего блока 131 коэффициентов преобразования. Другими словами, усиление α огибающей может определяться только на основе подмножества элементов 301 разрешения по частоте и/или только на основе подмножества полос 302 частот. В качестве примера усиление α огибающей может определяться на основе элементов 301 разрешения по частоте больше начального
35 элемента 304 разрешения по частоте (причем начальный элемент разрешения по частоте больше 0 или 1). В результате отрегулированная огибающая 139 для текущего блока 131 может определяться путем применения усиления α огибающей только к средним спектральным значениям 303 энергии у текущей интерполированной огибающей 136, которые ассоциируются с элементами 301 разрешения по частоте, находящимися выше
40 начального элемента 304 разрешения по частоте. Поэтому отрегулированная огибающая 139 для текущего блока 131 может соответствовать текущей интерполированной огибающей 136 для элементов 301 разрешения по частоте в начальном элементе разрешения по частоте и ниже его и может соответствовать текущей интерполированной огибающей 136, смещенной на усиление α огибающей, для элементов 301 разрешения по частоте выше начального элемента разрешения по частоте. Это иллюстрируется на
45 фиг.21а с помощью отрегулированной огибающей 339 (показанной пунктирными линиями).

Применение усиления α 137 огибающей (которое также называется усилением

коррекции уровня) к текущей интерполированной огибающей 136 соответствует регулировке или смещению текущей интерполированной огибающей 136, в силу этого приводя к отрегулированной огибающей 139, как проиллюстрировано с помощью фиг.21a. Усиление α 137 огибающей может кодироваться в виде данных 162 усиления

5 в потоке битов.

Кодер 100 может дополнительно содержать узел 107 уточнения огибающей, который конфигурируется для определения отрегулированной огибающей 139 на основе усиления α 137 огибающей и на основе текущей интерполированной огибающей 136.

Отрегулированную огибающую 139 можно использовать для обработки сигналов

10 блока 131 коэффициентов преобразования. Усиление α 137 огибающей можно квантовать до более высокого разрешения (например, с шагом в 1 дБ) по сравнению с текущей интерполированной огибающей 136 (которую можно квантовать с шагом в 3 дБ). По существу, отрегулированную огибающую 139 можно квантовать до более высокого разрешения усиления α 137 огибающей (например, с шагом в 1 дБ).

15 Кроме того, узел 107 уточнения огибающей может конфигурироваться для определения огибающей 138 распределения. Огибающая 138 распределения может соответствовать квантованной версии отрегулированной огибающей 139 (например, квантованной до уровней квантования в 3 дБ). Огибающая 138 распределения может использоваться с целью распределения битов. В частности, огибающая 138

20 распределения может использоваться для определения - для конкретного коэффициента преобразования текущего блока 131 - конкретного квантователя из заранее установленного набора квантователей, где конкретный квантователь должен использоваться для квантования конкретного коэффициента преобразования.

Кодер 100 содержит узел 108 выравнивания, сконфигурированный для выравнивания

25 текущего блока 131 с использованием отрегулированной огибающей 139, получая посредством этого блок 140 выровненных коэффициентов $\tilde{X}(k)$ преобразования. Блок 140 выровненных коэффициентов $\tilde{X}(k)$ преобразования может кодироваться с использованием контура предсказания в области преобразования. По существу, блок 140 может кодироваться с использованием блока 117 предсказания субполосы. Контур

30 предсказания содержит разностный узел 115, сконфигурированный для определения блока 141 коэффициентов $\Delta(k)$ ошибки предсказания на основе блока 140 выровненных коэффициентов $\tilde{X}(k)$ преобразования и на основе блока 150 предполагаемых коэффициентов $\hat{X}(k)$ преобразования, например $\Delta(k) = \tilde{X}(k) - \hat{X}(k)$. Следует отметить,

35 что вследствие того, что блок 140 содержит выровненные коэффициенты преобразования, то есть коэффициенты преобразования, которые нормализованы или выровнены с использованием значений 303 энергии отрегулированной огибающей 139, блок 150 предполагаемых коэффициентов преобразования также содержит оценки выровненных коэффициентов преобразования. Другими словами, разностный узел 115

40 работает в так называемой выровненной области. В результате блок 141 коэффициентов $\Delta(k)$ ошибки предсказания представляется в выровненной области.

Блок 141 коэффициентов $\Delta(k)$ ошибки предсказания может показывать дисперсию, которая отличается от единицы. Кодер 100 может содержать узел 111 изменения масштаба, сконфигурированный для изменения масштаба коэффициентов $\Delta(k)$ ошибки

45 предсказания, чтобы получить блок 142 коэффициентов ошибки с измененным масштабом. Узел 111 изменения масштаба может применять одно или несколько заранее установленных эвристических правил для выполнения изменения масштаба. В результате блок 142 коэффициентов ошибки с измененным масштабом показывает дисперсию,

которая (в среднем) ближе к единице (по сравнению с блоком 141 коэффициентов ошибки предсказания). Это может быть полезно для последующего квантования и кодирования.

5 Кодер 100 содержит узел 112 квантования коэффициентов, сконфигурированный для квантования блока 141 коэффициентов ошибки предсказания или блока 142
коэффициентов ошибки с измененным масштабом. Узел 112 квантования коэффициентов
может содержать или может применять набор заранее установленных квантователей.
Набор заранее установленных квантователей может предусматривать квантователи с
10 разными степенями точности или разным разрешением. Это иллюстрируется на фиг.22,
где иллюстрируются разные квантователи 321, 322, 323. Разные квантователи могут
обеспечивать разные уровни точности (указанные разными значениями в дБ).
Конкретный квантователь из множества квантователей 321, 322, 323 может
соответствовать конкретному значению огибающей 138 распределения. По существу
значение энергии огибающей 138 распределения может указывать на соответствующий
15 квантователь из множества квантователей. По существу, определение огибающей 138
распределения может упростить процесс выбора квантователя, используемого для
конкретного коэффициента ошибки. Другими словами, огибающая 138 распределения
может упростить процесс распределения битов.

Набор квантователей может содержать один или несколько квантователей 322,
20 которые применяют добавление псевдослучайного шума для рандомизации ошибки
квантования. Это иллюстрируется на фиг.22, показывающей первый набор 326 заранее
установленных квантователей, который содержит подмножество 324 квантователей с
добавлением псевдослучайного шума, и второй набор 327 заранее установленных
квантователей, который содержит подмножество 325 квантователей с добавлением
25 псевдослучайного шума. По существу, узел 112 квантования коэффициентов может
применять разные наборы 326, 327 заранее установленных квантователей, где набор
заранее установленных квантователей, который нужно использовать узлу 112
квантования коэффициентов, может зависеть от управляющего параметра 146,
предоставленного блоком предсказания 117 и/или определенного на основе другой
30 дополнительной информации, доступной в кодере и соответствующем декодере. В
частности, узел 112 квантования коэффициентов может конфигурироваться для выбора
набора 326, 327 заранее установленных квантователей для квантования блока 142
коэффициентов ошибки с измененным масштабом на основе управляющего параметра
146, где управляющий параметр 146 может зависеть от одного или нескольких
35 параметров блока предсказания, предоставленных блоком предсказания 117. Один или
несколько параметров блока предсказания может указывать качество блока 150
предполагаемых коэффициентов преобразования, предоставленного блоком
предсказания 117.

Квантованные коэффициенты ошибки могут энтропийно кодироваться с
40 использованием, например, кода Хаффмана, получая посредством этого данные 163
о коэффициентах, включаемые в поток битов, сформированный кодером 100.

Ниже описываются дополнительные подробности касательно выбора или определения
набора 326 квантователей 321, 322, 323. Набор 326 квантователей может соответствовать
упорядоченной совокупности 326 квантователей. Упорядоченная совокупность 326
45 квантователей может содержать N квантователей, где каждый квантователь может
соответствовать разному уровню искажения. По существу, совокупность 326
квантователей может обеспечивать N возможных уровней искажения. Квантователи в
совокупности 326 можно упорядочить в соответствии с уменьшающимся искажением

(или то же самое, что в соответствии с увеличивающимся SNR). Кроме того, квантователи можно обозначить целыми числами. В качестве примера квантователи можно обозначить 0, 1, 2, и т. п., где увеличивающееся целое число может указывать увеличивающееся SNR.

Совокупность 326 квантователей может быть такой, что интервал SNR между двумя последовательными квантователями постоянный, по крайней мере приблизительно. Например, SNR квантователя с обозначением "1" может составлять 1,5 дБ, а SNR квантователя с обозначением "2" может составлять 3,0 дБ. Поэтому квантователи в упорядоченной совокупности 326 квантователей могут быть такими, что при переходе от первого квантователя к соседнему второму квантователю SNR (отношение сигнал-шум) увеличивается практически на постоянное значение (например, 1,5 дБ) для всех пар из первого и второго квантователей.

Совокупность 326 квантователей может содержать:

- квантователь 321 с шумовым заполнением, который может обеспечить SNR, равное 0 дБ или немного ниже, которое для процесса распределения скорости может быть приблизительно равно 0 дБ;

- N_{dith} квантователей 322, которые могут использовать разностное добавление псевдослучайного шума и которые обычно соответствуют промежуточным уровням SNR (например, $N_{\text{dith}} > 0$); и

- N_{cq} классических квантователей 323, которые не используют разностное добавление псевдослучайного шума и которые обычно соответствуют сравнительно высоким уровням SNR (например, $N_{\text{cq}} > 0$). Квантователи 323 без добавления псевдослучайного шума могут соответствовать скалярным квантователям.

Общее количество N квантователей имеет вид $N = 1 + N_{\text{dith}} + N_{\text{cq}}$.

Пример совокупности 326 квантователей показан на фиг.24а. Квантователь 321 с шумовым заполнением из совокупности 326 квантователей можно реализовать, например, с использованием генератора случайных чисел, который выводит реализацию случайной переменной в соответствии с предопределенной статистической моделью.

К тому же совокупность 326 квантователей может содержать один или несколько квантователей 322 с добавлением псевдослучайного шума. Один или несколько квантователей с добавлением псевдослучайного шума могут формироваться с использованием реализации сигнала 602 с псевдослучайным шумом, как показано на фиг.24а. Сигнал 602 с псевдослучайным шумом может соответствовать блоку 602 значений псевдослучайного шума. Блок 602 псевдослучайных чисел может иметь такую же размерность, как размерность блока 142 коэффициентов ошибки с измененным масштабом, который нужно квантовать. Сигнал 602 с псевдослучайным шумом (или блок 602 значений псевдослучайного шума) может формироваться с использованием генератора 601 псевдослучайного шума. В частности, сигнал 602 с псевдослучайным шумом может формироваться с использованием справочной таблицы, содержащей равномерно распределенные случайные выборки.

Как будет показано применительно к фиг.24b, отдельные значения 632 псевдослучайного шума в блоке 602 значений псевдослучайного шума используются для применения псевдослучайного шума к соответствующему коэффициенту, который нужно квантовать (например, к соответствующему коэффициенту ошибки с измененным масштабом в блоке 142 коэффициентов ошибки с измененным масштабом). Блок 142 коэффициентов ошибки с измененным масштабом может содержать всего K коэффициентов ошибки с измененным масштабом. Аналогичным образом блок 602

значений псевдослучайного шума может содержать K значений 632 псевдослучайного шума. k^{oc} значение 632 псевдослучайного шума при $k = 1, \dots, K$ в блоке 602 значений псевдослучайного шума может применяться к $k^{ому}$ коэффициенту ошибки с измененным масштабом в блоке 142 коэффициентов ошибки с измененным масштабом.

Как указано выше, блок 602 значений псевдослучайного шума может иметь такое же измерение, как и блок 142 коэффициентов ошибки с измененным масштабом, которые нужно квантовать. Это выгодно, так как позволяет использовать один блок 602 значений псевдослучайного шума для всех квантователей 322 с добавлением псевдослучайного шума в совокупности 326 квантователей. Другими словами, чтобы квантовать и кодировать заданный блок 142 коэффициентов ошибки с измененным масштабом, псевдослучайный шум 602 может формироваться только один раз для всех допустимых совокупностей 326, 327 квантователей и для всех возможных распределений для искажения. Это облегчает достижение синхронности между кодером 100 и соответствующим декодером, так как использование одного сигнала 602 с псевдослучайным шумом не нужно явно сигнализировать соответствующему декодеру. В частности, кодер 100 и соответствующий декодер могут применять один и тот же генератор 601 псевдослучайного шума, который конфигурируется для формирования одного блока 602 значений псевдослучайного шума для блока 142 коэффициентов ошибки с измененным масштабом.

Состав совокупности 326 квантователей предпочтительно основывается на психоакустических соображениях. Низкоскоростное кодирование с преобразованием может приводить к спектральным артефактам, включая спектральные провалы и ограничение полосы, которые порождаются сущностью процесса "обратного заполнения водой" (water filling), который происходит в традиционных схемах квантования, которые применяются к коэффициентам преобразования. Слышимость спектральных провалов можно уменьшить путем введения шума в те полосы 302 частот, которые оказались ниже уровня воды за короткий период времени и которым, соответственно, была назначена нулевая скорость передачи битов.

Вообще, с помощью квантователя 322 с добавлением псевдослучайного шума можно добиться произвольно низкой скорости передачи битов. Например, в скалярном случае можно выбрать использование очень большого размера шага квантования. Тем не менее, операция с нулевой скоростью передачи битов на практике неосуществима, потому что она предъявляла бы требования к числовой точности, необходимой для работы квантователя с кодировщиком переменной длины. Это создает мотивацию к применению универсального квантователя 321 с шумовым заполнением к уровню искажения с SNR 0 дБ вместо применения квантователя 322 с добавлением псевдослучайного шума. Предложенная совокупность 326 квантователей проектируется так, что квантователи 322 с добавлением псевдослучайного шума используются для уровней искажения, которые ассоциируются с относительно небольшими размерами шага, так что кодирование переменной длины можно реализовать без необходимости решать проблемы, связанные с поддержанием числовой точности.

Для случая скалярного квантования квантователи 322 с разностным добавлением псевдослучайного шума можно реализовать с использованием последующих усилений, которые обеспечивают близкую к оптимальной производительность MSE. Пример скалярного квантователя 322 с разностным добавлением псевдослучайного шума показан на фиг.24b. Квантователь 322 с добавлением псевдослучайного шума содержит равномерный скалярный квантователь Q 612, который используется в структуре разностного добавления псевдослучайного шума. Структура разностного добавления

псевдослучайного шума содержит узел 611 вычитания псевдослучайного шума, который конфигурируется для вычитания значения 632 псевдослучайного шума (из блока 602 значений псевдослучайного шума) из соответствующего коэффициента ошибки (из блока 142 коэффициентов ошибки с измененным масштабом). Кроме того, структура разностного добавления псевдослучайного шума содержит соответствующий узел 613 добавления, который конфигурируется для добавления значения 632 псевдослучайного шума (из блока 602 значений псевдослучайного шума) к соответствующему скалярно квантованному коэффициенту ошибки. В проиллюстрированном примере узел 611 вычитания псевдослучайного шума размещается перед скалярным квантователем Q 612, а узел 613 добавления псевдослучайного шума размещается после скалярного квантователя Q 612. Значения 632 псевдослучайного шума из блока 602 значений псевдослучайного шума могут принимать значения из интервала $[-0,5, 0,5]$ или $[0, 1]$, умноженные на размер шага у скалярного квантователя 612. Следует отметить, что в альтернативной реализации квантователя 322 с добавлением псевдослучайного шума узел 611 вычитания псевдослучайного шума и узел 613 добавления псевдослучайного шума можно поменять друг с другом.

За структурой разностного добавления псевдослучайного шума может идти узел 614 масштабирования, который конфигурируется для изменения масштаба квантованных коэффициентов ошибки с помощью последующего усиления γ квантователя. После масштабирования квантованных коэффициентов ошибки получается блок 145 квантованных коэффициентов ошибки. Следует отметить, что вход X в квантователь 322 с добавлением псевдослучайного шума обычно соответствует коэффициентам в блоке 142 коэффициентов ошибки с измененным масштабом, которые попадают в конкретную полосу частот, которую нужно квантовать с использованием квантователя 322 с добавлением псевдослучайного шума. Аналогичным образом выход квантователя 322 с добавлением псевдослучайного шума обычно соответствует квантованным коэффициентам в блоке 145 квантованных коэффициентов ошибки, которые попадают в конкретную полосу частот.

Можно предположить, что вход X в квантователь 322 с добавлением псевдослучайного шума является нулевым средним, и что известна дисперсия $\sigma_X^2 = E\{X^2\}$ входа X. (Например, дисперсия сигнала может определяться из огибающей этого сигнала). Кроме того, можно предположить, что блок Z 602 псевдослучайного шума, содержащий значения 632 псевдослучайного шума, доступен кодеру 100 и соответствующему декодеру. Кроме того, можно предположить, что значения 632 псевдослучайного шума не зависят от входа X. Можно использовать различные другие псевдослучайные шумы 602, но в дальнейшем предполагается, что псевдослучайный шум Z 602 равномерно распределен между 0 и Δ , что можно обозначить с помощью U $(0, \Delta)$. На практике можно использовать любой псевдослучайный шум, который выполняет так называемые условия Шухмана (например, псевдослучайный шум 602, который равномерно распределен между $[-0,5, 0,5]$, умноженное на размер Δ шага скалярного квантователя 612).

Квантователь Q 612 может быть решеткой, и размером ее ячейки Вороного может быть Δ . В этом случае сигнал с псевдослучайным шумом имел бы равномерное распределение по размеру ячейки Вороного в используемой решетке.

Последующее усиление γ квантователя можно вывести, принимая во внимание дисперсию сигнала и размер шага квантования, поскольку квантователь с добавлением псевдослучайного шума является аналитически определяемым для любого размера шага (то есть скорости передачи битов). В частности, можно вывести последующее

усиление для повышения производительности MSE у квантователя с разностным добавлением псевдослучайного шума. Последующее усиление может иметь вид:

$$\gamma = \frac{\sigma_x^2}{\sigma_x^2 + \frac{\Delta^2}{12}}.$$

Даже если путем применения последующего усиления γ можно повысить производительность MSE у квантователя 322 с добавлением псевдослучайного шума, квантователь 322 с добавлением псевдослучайного шума обычно обладает меньшей производительностью MSE, нежели квантователь без добавления псевдослучайного шума (хотя эта потеря производительности устраняется, когда увеличивается скорость передачи битов). Следовательно, квантователи с добавлением псевдослучайного шума обычно более шумные, чем их версии без добавления псевдослучайного шума. Поэтому может быть желательно использовать квантователи 322 с добавлением псевдослучайного шума только тогда, когда использование квантователей 322 с добавлением псевдослучайного шума оправдано благоприятным для восприятия свойством шумового заполнения у квантователей 322 с добавлением псевдослучайного шума.

Поэтому можно предоставить совокупность 326 квантователей, содержащую три типа квантователей. Упорядоченная совокупность 326 квантователей может содержать один квантователь 321 с шумовым заполнением, один или несколько квантователей 322 с разностным добавлением псевдослучайного шума и один или несколько классических (без добавления псевдослучайного шума) квантователей 323. Последовательные квантователи 321, 322, 323 могут обеспечить постепенные улучшения в SNR. Постепенные улучшения между парой соседних квантователей в упорядоченной совокупности 326 квантователей могут быть практически постоянными для некоторых или всех пар соседних квантователей.

Конкретная совокупность 326 квантователей может задаваться количеством квантователей 322 с добавлением псевдослучайного шума и количеством квантователей 323 без добавления псевдослучайного шума, содержащимися в конкретной совокупности 326. Кроме того, конкретная совокупность 326 квантователей может задаваться конкретной реализацией сигнала 602 с псевдослучайным шумом. Совокупность 326 может проектироваться для того, чтобы обеспечивать эффективное для восприятия квантование представления коэффициентов преобразования: шумовое заполнение при нулевой скорости (дающее SNR, равное 0 дБ или немного ниже); шумовое заполнение с помощью разностного добавления псевдослучайного шума на промежуточном уровне искажения (промежуточное SNR); и отсутствие шумового заполнения на низких уровнях искажения (высокое SNR). Совокупность 326 предоставляет набор допустимых квантователей, которые могут выбираться во время процесса распределения скорости. Применение конкретного квантователя из совокупности 326 квантователей к коэффициентам конкретной полосы 302 частот определяется во время процесса распределения скорости. Обычно заранее не известно, какой квантователь будет использоваться для квантования коэффициентов конкретной полосы 302 частот. Однако обычно заранее известно, каков состав совокупности 326 квантователей.

Аспект использования разных типов квантователей для разных полос 302 частот в блоке 142 коэффициентов ошибки иллюстрируется на фиг.24с, где показан примерный результат процесса распределения скорости. В этом примере предполагается, что распределение скорости придерживается так называемого принципа "обратного заполнения водой". Фиг.24с иллюстрирует спектр 625 входного сигнала (или огибающую квантуемого блока коэффициентов). Видно, что полоса 623 частот обладает

относительно большой спектральной энергией и квантуется с использованием классического квантователя 323, который обеспечивает сравнительно низкие уровни искажения. Полосы 622 частот показывают спектральную энергию выше уровня 624 воды. Коэффициенты в этих полосах 622 частот могут квантоваться с использованием квантователей 322 с добавлением псевдослучайного шума, которые обеспечивают промежуточные уровни искажения. Полосы 621 частот показывают спектральную энергию ниже уровня 624 воды. Коэффициенты в этих полосах 621 частот могут квантоваться с использованием шумового заполнения при нулевой скорости. Разные квантователи, используемые для квантования конкретного блока коэффициентов (представленного спектром 625), могут быть частью конкретной совокупности 326 квантователей, которая определена для конкретного блока коэффициентов.

Поэтому три разных типа квантователей 321, 322, 323 могут применяться выборочно (например, выборочно в отношении частоты). Решение о применении конкретного типа квантователя может приниматься применительно к процедуре распределения скорости, которая описывается ниже. Процедура распределения скорости может применять критерий восприятия, который можно вывести из огибающей RMS входного сигнала (или, например, из спектральной плотности мощности сигнала). Тип квантователя для применения в конкретной полосе 302 частот не нужно сигнализировать соответствующему декодеру явно. Необходимость сигнализации выбранного типа квантователя устраняется, поскольку соответствующий декодер способен определить конкретный набор 326 квантователей, который использовался для квантования блока входного сигнала, из лежащего в основе критерия восприятия (например, огибающей 138 распределения), из заранее установленного состава совокупности квантователей (например, заранее установленного набора разных совокупностей квантователей) и из одного параметра глобального распределения скорости (также называемого параметром смещения).

Определение совокупности 326 квантователей, которая использована кодером 100, в декодере упрощается путем проектирования совокупности 326 квантователей так, чтобы квантователи упорядочивались в соответствии с их искажением (например, SNR). Каждый квантователь в совокупности 326 может уменьшить искажение (может улучшить SNR) предыдущего квантователя на постоянное значение. Кроме того, конкретная совокупность 326 квантователей может ассоциироваться с одной реализацией псевдослучайного сигнала 602 с псевдослучайным шумом в течение всего процесса распределения скорости. Вследствие этого результат процедуры распределения скорости не влияет на реализацию сигнала 602 с псевдослучайным шумом. Это выгодно для обеспечения сходимости процедуры распределения скорости. Кроме того, это дает декодеру возможность выполнять декодирование, если декодеру известна одна реализация сигнала 602 с псевдослучайным шумом. Декодеру можно сообщить реализацию сигнала 602 с псевдослучайным шумом с использованием одинакового генератора 601 псевдослучайного шума в кодере 100 и в соответствующем декодере.

Как указано выше, кодер 100 может конфигурироваться для выполнения процесса распределения битов. С этой целью кодер 100 может содержать узлы 109, 110 распределения битов. Узел 109 распределения битов может конфигурироваться для определения общего количества 143 битов, которые доступны для кодирования текущего блока 142 коэффициентов ошибки с измененным масштабом. Общее количество 143 битов может определяться на основе огибающей 138 распределения. Узел 110 распределения битов может конфигурироваться для предоставления относительного распределения битов разным коэффициентам ошибки с измененным масштабом в

зависимости от соответствующего значения энергии в огибающей 138 распределения.

Процесс распределения битов может применять процедуру итеративного распределения. В ходе процедуры распределения огибающую 138 распределения можно смещать с использованием параметра смещения, посредством этого выбирая
 5 квантователи с увеличенным/уменьшенным разрешением. По существу, параметр смещения можно использовать для уточнения или огрубления общего квантования. Параметр смещения может определяться так, что данные 163 о коэффициентах, которые
 10 получают с использованием квантователей, заданных параметром смещения и огибающей 138 распределения, содержат количество битов, которое соответствует общему количеству 143 битов, назначенных текущему блоку 131 (или не превышает его). Параметр смещения, который использован кодером 100 для кодирования текущего
 15 блока 131, включается в виде данных 163 о коэффициентах в поток битов. В результате соответствующему декодеру предоставляется возможность определить квантователи, которые использованы узлом 112 квантования коэффициентов для квантования блока
 142 коэффициентов ошибки с измененным масштабом.

По существу процесс распределения скорости может выполняться в кодере 100, где он стремится распределить доступные биты 143 в соответствии с моделью восприятия. Модель восприятия может зависеть от огибающей 138 распределения, выведенной из
 20 блока 131 коэффициентов преобразования. Алгоритм распределения скорости распределяет доступные биты 143 между разными типами квантователей, то есть с шумовым заполнением 321 при нулевой скорости, одним или несколькими квантователями 322 с добавлением псевдослучайного шума и одним или несколькими классическими квантователями 323 без добавления псевдослучайного шума.
 Окончательное решение о типе квантователя, используемого для квантования
 25 коэффициентов конкретной полосы 302 частот спектра, может зависеть от модели восприятия сигнала, от реализации псевдослучайного шума и от ограничения скорости передачи битов.

В соответствующем декодере распределение битов (указанное огибающей 138 распределения и параметром смещения) можно использовать для определения
 30 вероятностей индексов квантования, чтобы упростить декодирование без потерь. Можно использовать способ вычисления вероятностей индексов квантования, который применяет реализацию псевдослучайного шума 602 полной полосы, модель восприятия, параметризованную огибающей 138 сигнала и параметром распределения скорости
 (то есть параметром смещения). При использовании огибающей 138 распределения,
 35 параметра смещения и сведений о блоке 602 значений псевдослучайного шума состав совокупности 326 квантователей в декодере может соответствовать совокупности 326, используемой в кодере 100.

Как указывалось выше, ограничение скорости передачи битов может задаваться в виде максимального разрешенного количества 143 битов на кадр. Это применяется,
 40 например, к индексам квантования, которые впоследствии энтропийно кодируются с использованием, например, кода Хаффмана. В частности, это применяется в сценариях кодирования, где поток битов формируется последовательно, где одновременно квантуется один параметр, и где соответствующий индекс квантования преобразуется в двоичное кодовое слово, которое добавляется к потоку битов.

Если используется арифметическое кодирование (или кодирование диапазона), то принцип отличается. Обычно одно кодовое слово назначается длинной
 45 последовательности индексов квантования, применительно к арифметическому кодированию. Обычно нельзя ассоциировать точно конкретную часть потока битов с

конкретным параметром. В частности, применительно к арифметическому кодированию обычно неизвестно количество битов, которое необходимо для кодирования случайной реализации сигнала. Это имеет место, даже если известна статистическая модель сигнала.

Чтобы решить вышеупомянутую техническую проблему, предлагается сделать арифметический кодер частью алгоритма распределения скорости. Во время процесса распределения скорости кодер пытается квантовать и кодировать набор коэффициентов одной или нескольких полос 302 частот. Для каждой такой попытки можно наблюдать изменение состояния арифметического кодера и вычислять количество положений для продвижения в потоке битов (вместо вычисления количества битов). Если устанавливается ограничение максимальной скорости передачи битов, то это ограничение максимальной скорости передачи битов можно использовать в процедуре распределения скорости. Стоимость битов завершения арифметического кода может включаться в стоимость последнего кодированного параметра, и обычно стоимость битов завершения меняется в зависимости от состояния арифметического кодировщика. Тем не менее, как только доступна стоимость завершения, можно определить количество битов, необходимое для кодирования индексов квантования, соответствующих набору коэффициентов одной или нескольких полос 302 частот.

Следует отметить, что применительно к арифметическому кодированию можно использовать одну реализацию псевдослучайного шума 602 для всего процесса распределения скорости (в конкретном блоке 142 коэффициентов). Как указывалось выше, арифметический кодер можно использовать для оценки стоимости скорости передачи битов у выбора конкретного квантователя в процедуре распределения скорости. Можно наблюдать изменение состояния арифметического кодера, и изменение состояния может использоваться для вычисления количества битов, необходимого для выполнения квантования. Кроме того, в процессе распределения скорости можно использовать процесс завершения арифметического кода.

Как указано выше, индексы квантования могут кодироваться с использованием арифметического кода или энтропийного кода. Если индексы квантования кодируются энтропийно, то можно принять во внимание распределение вероятностей у индексов квантования, чтобы назначить кодовые слова переменной длины отдельному индексу или группам индексов квантования. Использование добавления псевдослучайного шума может влиять на распределение вероятностей у индексов квантования. В частности, конкретная реализация сигнала 602 с псевдослучайным шумом может влиять на распределение вероятностей у индексов квантования. Вследствие практически неограниченного количества реализаций сигнала 602 с псевдослучайным шумом в общем случае вероятности кодовых слов заранее неизвестны, и невозможно использовать кодирование методом Хаффмана.

Авторы изобретения обратили внимание, что можно уменьшить количество возможных реализаций псевдослучайного шума до относительно небольшого и управляемого набора реализаций сигнала 602 с псевдослучайным шумом. В качестве примера ограниченный набор значений псевдослучайного шума можно предоставить для каждой полосы 302 частот. С этой целью кодер 100 (а также соответствующий декодер) может содержать дискретный генератор 801 псевдослучайного шума, сконфигурированный для формирования сигнала 602 с псевдослучайным шумом путем выбора одной из M заранее установленных реализаций псевдослучайного шума (см. фиг.26). В качестве примера для каждой полосы 302 частот можно использовать M разных заранее установленных реализаций псевдослучайного шума. Количество M заранее установленных реализаций псевдослучайного шума может быть $M < 5$ (например,

$M = 4$ или $M = 3$).

Благодаря ограниченному количеству M реализаций псевдослучайного шума можно подготовить кодовую книгу Хаффмана (по возможности многомерную) для каждой реализации псевдослучайного шума, получая в результате совокупность 803 из M кодовых книг. Кодер 100 может содержать узел 802 выбора кодовой книги, который конфигурируется для выбора одной из совокупности 803 из M заранее установленных кодовых книг на основе выбранной реализации псевдослучайного шума. При этом обеспечивается, что энтропийное кодирование синхронно с формированием псевдослучайного шума. Выбранная кодовая книга 811 может использоваться для кодирования отдельного индекса или групп индексов квантования, которые квантованы с использованием выбранной реализации псевдослучайного шума. В результате можно повысить производительность энтропийного кодирования при использовании квантователей с добавлением псевдослучайного шума.

Совокупность 803 заранее установленных кодовых книг и дискретный генератор 801 псевдослучайного шума также могут использоваться в соответствующем декодере (как проиллюстрировано на фиг.26). Декодирование осуществимо, если используется псевдослучайный шум, и если декодер остается синхронным с кодером 100. В этом случае дискретный генератор 801 псевдослучайного шума в декодере формирует сигнал 602 с псевдослучайным шумом, и конкретная реализация псевдослучайного шума однозначно ассоциируется с конкретной кодовой книгой 811 Хаффмана из совокупности 803 кодовых книг. Принимая во внимание психоакустическую модель (например, представленную огибающей 138 распределения и параметром распределения скорости) и выбранную кодовую книгу 811, декодер может выполнить декодирование с использованием декодера Хаффмана 551, чтобы получить декодированные индексы 812 квантования.

По существу, вместо арифметического кодирования можно использовать сравнительно небольшой набор 803 кодовых книг Хаффмана. Использование конкретной кодовой книги 811 из набора 813 кодовых книг Хаффмана может зависеть от заранее установленной реализации сигнала 602 с псевдослучайным шумом. Вместе с тем можно использовать ограниченный набор допустимых значений псевдослучайного шума, образующий M заранее установленных реализаций псевдослучайного шума. Процесс распределения скорости тогда может включать в себя использование квантователей без добавления псевдослучайного шума, квантователей с добавлением псевдослучайного шума и кодирование методом Хаффмана.

В результате квантования коэффициентов ошибки с измененным масштабом получается блок 145 квантованных коэффициентов ошибки. Блок 145 квантованных коэффициентов ошибки соответствует блоку коэффициентов ошибки, которые доступны в соответствующем декодере. Следовательно, блок 145 квантованных коэффициентов ошибки можно использовать для определения блока 150 предполагаемых коэффициентов преобразования. Кодер 100 может содержать узел 113 обратного изменения масштаба, сконфигурированный для выполнения инверсии к операциям изменения масштаба, выполняемым узлом 113 изменения масштаба, посредством этого получая блок 147 масштабированных квантованных коэффициентов ошибки. Узел 116 добавления можно использовать для определения блока 148 восстановленных выровненных коэффициентов путем добавления блока 150 предполагаемых коэффициентов преобразования к блоку 147 масштабированных квантованных коэффициентов ошибки. Кроме того, узел 114 обратного выравнивания может использоваться для применения отрегулированной огибающей 139 к блоку 148 восстановленных выровненных коэффициентов, посредством

этого получая блок 149 восстановленных коэффициентов. Блок 149 восстановленных коэффициентов соответствует версии блока 131 коэффициентов преобразования, которая доступна в соответствующем декодере. В результате блок 149 восстановленных коэффициентов может использоваться в блоке предсказания 117 для определения блока 150 предполагаемых коэффициентов.

Блок 149 восстановленных коэффициентов представляется в невыровненной области, то есть блок 149 восстановленных коэффициентов также представляет огибающую спектра текущего блока 131. Как указано ниже, это может быть полезно для производительности блока 117 предсказания.

Блок предсказания 117 может конфигурироваться для оценивания блока 150 предполагаемых коэффициентов преобразования на основе одного или нескольких предыдущих блоков 149 восстановленных коэффициентов. В частности, блок предсказания 117 может конфигурироваться для определения одного или нескольких параметров блока предсказания так, что уменьшается заранее установленный критерий ошибок предсказания (например, минимизируется). В качестве примера один или несколько параметров блока предсказания могут определяться так, что уменьшается энергия, или взвешенная по восприятию энергия, блока 141 коэффициентов ошибки предсказания (например, минимизируется). Один или несколько параметров блока предсказания могут включаться в виде данных 164 блока предсказания в поток битов, сформированный кодером 100.

Блок предсказания 117 может применять модель прохождения сигналов, которая описана в заявке на патент US61750052 и заявках на патент, которые притязают на ее приоритет, содержимое которых включается в этот документ посредством ссылки. Один или несколько параметров блока предсказания могут соответствовать одному или нескольким параметрам модели у модели прохождения сигналов.

Фиг.19b показывает блок-схему дополнительного примерного речевого кодера 170 с преобразованием. Речевой кодер 170 с преобразованием из фиг.19b содержит многие компоненты кодера 100 из фиг.19a. Однако речевой кодер 170 с преобразованием из фиг.19b конфигурируется для формирования потока битов, имеющего переменную скорость передачи битов. С этой целью кодер 170 содержит узел 172 состояния средней скорости передачи битов (ABR), сконфигурированный для отслеживания скорости передачи битов, которая использована потоком битов для предыдущих блоков 131. Узел 171 распределения битов использует эту информацию для определения общего количества 143 битов, которое доступно для кодирования текущего блока 131 коэффициентов преобразования.

Ниже соответствующий речевой декодер 500 с преобразованием описывается применительно к фиг.23a-23d. Фиг.23a показывает блок-схему примерного речевого декодера 500 с преобразованием. Блок-схема показывает гребенку 504 фильтров синтеза (также называемую узлом обратного преобразования), которая используется для преобразования блока 149 восстановленных коэффициентов из области преобразования во временную область, посредством этого получая выборки декодированного аудиосигнала. Гребенка 504 фильтров синтеза может применять обратное MDCT с заранее установленным шагом (например, с шагом приблизительно в 5 мс или 256 выборок).

Основной контур декодера 500 работает в единицах этого шага. Каждый этап создает вектор области преобразования (также называемый блоком), имеющий длину или измерение, которое соответствует заранее установленной полосе пропускания системы. После заполнения нулями до размера преобразования у гребенки 504 фильтров синтеза

вектор области преобразования будет использоваться для синтеза обновления сигнала временной области с заранее установленной длиной (например, 5 мс) для процесса перекрытия/добавления в гребенке 504 фильтров синтеза.

Как указано выше, универсальные аудиокодеки с преобразованием обычно применяют кадры с последовательностями коротких блоков в диапазоне 5 мс для обработки переходов. По существу, универсальные аудиокодеки с преобразованием предоставляют необходимые преобразования и инструменты переключения кадров для непрерывного совместного существования коротких и длинных блоков. Поэтому речевую спектральную предварительную обработку, заданную путем исключения гребенки 504 фильтров синтеза из фиг.23а, можно легко интегрировать в универсальный аудиокодек с преобразованием без необходимости внедрения дополнительных инструментов переключения. Другими словами, речевой декодер 500 с преобразованием из фиг.23а можно легко объединить с универсальным аудиодекодером с преобразованием. В частности, речевой декодер 500 с преобразованием из фиг.23а может применять гребенку 504 фильтров синтеза, предоставленную универсальным аудиодекодером с преобразованием (например, декодером AAC или HE-AAC).

Из входящего потока битов (в частности, из данных 161 огибающей и из данных 162 усиления, содержащихся в потоке битов) декодер 503 огибающей может определить огибающую сигнала. В частности, декодер 503 огибающей может конфигурироваться для определения отрегулированной огибающей 139 на основе данных 161 огибающей и данных 162 усиления. По существу, декодер 503 огибающей может выполнять задачи, аналогичные узлу 104 интерполяции и узлу 107 уточнения огибающей в кодере 100, 170. Как указывалось выше, отрегулированная огибающая 109 представляет модель дисперсии сигнала в наборе предопределенных полос 302 частот.

Кроме того, декодер 500 содержит узел 114 обратного выравнивания, который конфигурируется для применения отрегулированной огибающей 139 к вектору выровненной области, чьи элементы номинально могут иметь единичную дисперсию. Вектор выровненной области соответствует блоку 148 восстановленных выровненных коэффициентов, описанному применительно к кодери 100, 170. На выходе узла 114 обратного выравнивания получается блок 149 восстановленных коэффициентов. Блок 149 восстановленных коэффициентов предоставляется в гребенку 504 фильтров синтеза (для формирования декодированного аудиосигнала) и в блок предсказания 517 субполосы.

Блок предсказания 517 субполосы работает аналогично блоку предсказания 117 в кодере 100, 170. В частности, блок предсказания 517 субполосы конфигурируется для определения блока 150 предполагаемых коэффициентов преобразования (в выровненной области) на основе одного или нескольких предыдущих блоков 149 восстановленных коэффициентов (используя один или несколько параметров блока предсказания, сигнализированных в потоке битов). Другими словами, блок предсказания 517 субполосы конфигурируется для вывода предсказанного вектора выровненной области из буфера ранее декодированных выходных векторов и огибающих сигнала на основе параметров блока предсказания, например отставания блока предсказания и усиления блока предсказания. Декодер 500 содержит декодер 501 блока предсказания, сконфигурированный для декодирования данных 164 блока предсказания, чтобы определить один или несколько параметров блока предсказания.

Декодер 500 дополнительно содержит декодер 502 спектра, который конфигурируется для внесения аддитивной поправки в предсказанный вектор выровненной области, обычно на основе наибольшей части потока битов (то есть на основе данных 163 о

коэффициентах). Процесс декодирования спектра управляется преимущественно вектором распределения, который выводится из огибающей и переданного управляющего параметра распределения (также называемого параметром смещения). Как проиллюстрировано на фиг.23а, может иметь место прямая зависимость декодера 502 спектра от параметров 520 блока предсказания. По существу, декодер 502 спектра может конфигурироваться для определения блока 147 масштабированных квантованных коэффициентов ошибки на основе принятых данных 163 о коэффициентах. Как указано применительно к кодеру 100, 170, квантователи 321, 322, 323, используемые для квантования блока 142 коэффициентов ошибки с измененным масштабом, обычно

зависят от огибающей 138 распределения (которую можно вывести из отрегулированной огибающей 139) и от параметра смещения. Кроме того, квантователи 321, 322, 323 могут зависеть от управляющего параметра 146, предоставленного блоком предсказания 117. Управляющий параметр 146 можно вывести с помощью декодера 500, используя параметры 520 блока предсказания (аналогично кодеру 100, 170).

Как указано выше, принятый поток битов содержит данные 161 огибающей и данные 162 усиления, которые можно использовать для определения отрегулированной огибающей 139. В частности, узел 531 в декодере 503 огибающей может конфигурироваться для определения квантованной текущей огибающей 134 из данных 161 огибающей. В качестве примера квантованная текущая огибающая 134 может иметь разрешение 3 дБ в predetermined полосах 302 частот (как указано на фиг.21а). Квантованная текущая огибающая 134 может обновляться для каждого набора 132, 332 блоков (например, каждые четыре единицы кодирования, то есть блока, или каждые 20 мс), в частности, для каждого сдвинутого набора 332 блоков. Полосы 302 частот у квантованной текущей огибающей 134 могут содержать увеличивающееся количество элементов 301 разрешения по частоте в зависимости от частоты, чтобы приспособиться к свойствам слуха человека.

Квантованную текущую огибающую 134 можно линейно интерполировать от квантованной предыдущей огибающей 135 в интерполированные огибающие 136 для каждого блока 131 в сдвинутом наборе 332 блоков (или, возможно, в текущем наборе 132 блоков). Интерполированные огибающие 136 могут определяться в квантованной области 3 дБ. Это означает, что интерполированные значения 303 энергии могут округляться до ближайшего уровня 3 дБ. Примерная интерполированная огибающая 136 иллюстрируется пунктирным графиком на фиг.21а. Для каждой квантованной текущей огибающей 134 предоставляются четыре усиления α 137 коррекции уровня (также называемые усилениями огибающей) в виде данных 162 усиления. Узел 532 декодирования усиления может конфигурироваться для определения усиления α 137 коррекции уровня из данных 162 усиления. Усиления коррекции уровня можно квантовать с шагом в 1 дБ. Каждое усиление коррекции уровня применяется к соответствующей интерполированной огибающей 136, чтобы предоставить отрегулированные огибающие 139 для разных блоков 131. Благодаря повышенному разрешению усиления 137 коррекции уровня отрегулированная огибающая 139 может иметь повышенное разрешение (например, разрешение в 1 дБ).

Фиг.21b показывает примерную линейную или геометрическую интерполяцию между квантованной предыдущей огибающей 135 и квантованной текущей огибающей 134. Огибающие 135, 134 можно разделить на часть среднего уровня и часть формы логарифмического спектра. Эти части можно интерполировать с помощью независимых стратегий, например линейной, геометрической или гармонической (параллельные резисторы) стратегии. По существу, можно использовать разные схемы интерполяции

для определения интерполированных огибающих 136. Схема интерполяции, используемая декодером 500, обычно соответствует схеме интерполяции, используемой кодером 100, 170.

Узел 107 уточнения огибающей в декодере 503 огибающей может конфигурироваться для определения огибающей 138 распределения из отрегулированной огибающей 139 путем квантования отрегулированной огибающей 139 (например, шагами в 3 дБ). Огибающая 138 распределения может использоваться в сочетании с управляющим параметром распределения или параметром смещения (содержащимся в данных 163 о коэффициентах) для создания номинального целочисленного вектора распределения, используемого для управления спектральным декодированием, то есть декодированием данных 163 о коэффициентах. В частности, номинальный целочисленный вектор распределения может использоваться для определения квантователя для обратного квантования индексов квантования, содержащихся в данных 163 о коэффициентах. Огибающая 138 распределения и номинальный целочисленный вектор распределения могут определяться в кодере 100, 170 и в декодере 500 аналогичным образом.

Фиг.27 иллюстрирует примерный процесс распределения битов на основе огибающей 138 распределения. Как указывалось выше, огибающую 138 распределения можно квантовать в соответствии с заранее установленным разрешением (например, разрешением в 3 дБ). Каждое квантованное спектральное значение энергии огибающей 138 распределения можно назначить соответствующему целому значению, где соседние целые значения могут представлять разность спектральной энергии, соответствующую заранее установленному разрешению (например, разность 3 дБ). Результирующий набор целых чисел может называться целочисленной огибающей 1004 распределения (называемой *iEnv*). Целочисленную огибающую 1004 распределения можно сместить на параметр смещения, чтобы получить номинальный целочисленный вектор распределения (называемый *iAlloc*), который обеспечивает прямое указание квантователя, используемого для квантования коэффициента конкретной полосы 302 частот (идентифицированной индексом полосы частот, *bandIdx*).

Фиг.27 показывает на схеме 1003 целочисленную огибающую 1004 распределения в зависимости от полос 302 частот. Видно, что для полосы 1002 частот (*bandIdx* = 7) целочисленная огибающая 1004 распределения принимает целое значение -17 (*iEnv*[7] = -17). Целочисленная огибающая 1004 распределения может быть ограничена максимальным значением (называемым *iMax*, например *iMax* = -15). Процесс распределения битов может применять формулу распределения битов, которая предоставляет индекс 1006 квантователя (называемый *iAlloc* [*bandIdx*]) в зависимости от целочисленной огибающей 1004 распределения и параметра смещения (называемого *AllocOffset*). Как указывалось выше, параметр смещения (то есть *AllocOffset*) передается соответствующему декодеру 500, посредством этого предоставляя декодеру 500 возможность определить индексы 1006 квантователей с использованием формулы распределения битов. Формула распределения битов может иметь вид:

$$iAlloc[bandIdx] = iEnv[bandIdx] - (iMax - CONSTANT_OFFSET) + AllocOffset,$$

где *CONSTANT_OFFSET* может быть постоянным смещением, например *CONSTANT_OFFSET* = 20. В качестве примера, если процесс распределения битов определил, что ограничения скорости передачи битов можно добиться с использованием параметра смещения *AllocOffset* = -13, то индекс 1007 квантователя у 7^{ой} полосы частот можно получить в виде *iAlloc*[7] = -17 - (-15-20) - 13 = 5. С помощью использования вышеупомянутой формулы распределения битов для всех полос 302 частот можно определить индексы 1006 квантователей (и в результате, квантователи 321, 322, 323)

для всех полос 302 частот. Индекс квантователя меньше нуля можно округлить в большую сторону до нулевого индекса квантователя. Аналогичным образом индекс квантователя больше максимального доступного индекса квантователя можно округлить в меньшую сторону до максимального доступного индекса квантователя.

5 Кроме того, фиг.27 показывает примерную огибающую 1011 шума, которую можно достичь с использованием схемы квантования, описанной в настоящем документе. Огибающая 1011 шума показывает огибающую шума квантования, который вносится во время квантования. При изображении вместе с огибающей сигнала (представленной целочисленной огибающей 1004 распределения на фиг.27) огибающая 1011 шума иллюстрирует, что распределение шума квантования оптимизировано для восприятия относительно огибающей сигнала.

Чтобы позволить декодеру 500 синхронизироваться с принимаемым потоком битов, могут передаваться разные типы кадров. Кадр может соответствовать набору 132, 332 блоков, в частности, сдвинутому блоку 332 блоков. В частности, могут передаваться так называемые Р-кадры, которые кодируются по отношению к предыдущему кадру. В вышеприведенном описании допускалось, что декодер 500 знает о квантованной предыдущей огибающей 135. Квантованная предыдущая огибающая 135 может предоставляться в предыдущем кадре, так что текущий набор 132 или соответствующий сдвинутый набор 332 может соответствовать Р-кадру. Однако в сценарии запуска декодер 500 обычно не знает о квантованной предыдущей огибающей 135. С этой целью может передаваться I-кадр (например, при запуске или систематически). I-кадр может содержать две огибающие, одна из которых используется в качестве квантованной предыдущей огибающей 135, а другая используется в качестве квантованной текущей огибающей 134. I-кадры могут использоваться для случая запуска речевой спектральной предварительной обработки (то есть речевого декодера 500 с преобразованием), например, после кадра, применяющего другой режим аудиокодирования, и/или в качестве инструмента для явного разрешения точки склейки аудиопотока битов.

Работа блока 517 предсказания субполосы иллюстрируется на фиг.23d. В проиллюстрированном примере параметрами 520 блока предсказания являются параметр отставания и параметр g усиления блока предсказания. Параметры 520 блока предсказания могут определяться из данных 164 блока предсказания с использованием заранее установленной таблицы возможных значений для параметра отставания и параметра усиления блока предсказания. Это дает возможность эффективной по скорости передачи параметров 520 блока предсказания.

35 Один или несколько ранее декодированных векторов коэффициентов преобразования (то есть один или несколько предыдущих блоков 149 восстановленных коэффициентов) можно сохранить в буфере 541 сигнала субполосы (или MDCT). Буфер 541 может обновляться в соответствии с шагом (например, каждые 5 мс). Выделитель 543 блока предсказания может конфигурироваться для воздействия на буфер 541 в зависимости от нормализованного параметра T отставания. Нормализованный параметр T отставания может определяться путем нормализации параметра 520 отставания до единиц шагов (например, до единиц шагов MDCT). Если параметр T отставания является целым числом, то выделитель 543 может отобразить в буфер 541 T единиц времени одного или нескольких ранее декодированных векторов коэффициентов преобразования.

45 Другими словами, параметр T отставания может указывать, какие из одного или нескольких предыдущих блоков 149 восстановленных коэффициентов нужно использовать для определения блока 150 предполагаемых коэффициентов преобразования. Подробное обсуждение касательно возможной реализации выделителя

543 предоставляется в заявке на патент US61750052 и заявках на патент, которые притязают на ее приоритет, содержимое которых включается в этот документ посредством ссылки.

Выделитель 543 может воздействовать на векторы (или блоки), несущие огибающие
 5 полного сигнала. С другой стороны, блок 150 предполагаемых коэффициентов преобразования (предоставляемый блоком предсказания 517 субполосы) представляется в выровненной области. Следовательно, выход выделителя 543 можно оформить в виде вектора выровненной области. Это может достигаться с использованием формирователя 544, который применяет отрегулированные огибающие 139 из одного или нескольких
 10 предыдущих блоков 149 восстановленных коэффициентов. Отрегулированные огибающие 139 из одного или нескольких предыдущих блоков 149 восстановленных коэффициентов можно сохранить в буфере 542 огибающих. Узел 544 формирователя может конфигурироваться для выборки в буфер 542 огибающих огибающей задержанного сигнала, используемой при выравнивании, из T_0 единиц времени, где T_0
 15 - ближайшее к T целое число. Тогда вектор выровненной области можно масштабировать с помощью параметра g усиления, чтобы получить блок 150 предполагаемых коэффициентов преобразования (в выровненной области).

В качестве альтернативы процесс задержанного выравнивания, выполняемый формирователем 544, можно пропустить при использовании блока 517 предсказания
 20 субполосы, который работает в выровненной области, например блока 517 предсказания субполосы, который воздействует на блоки 148 восстановленных выровненных коэффициентов. Однако обнаружено, что последовательность векторов выровненной области (или блоков) не отображается правильно в сигналы времени из-за наложенных по времени аспектов преобразования (например, преобразования MDCT). В результате
 25 у выделителя 543 уменьшается соответствие лежащей в основе модели прохождения сигналов, и от альтернативной структуры получается более высокий уровень шумов кодирования. Другими словами, обнаружено, что используемые блоком предсказания 517 субполосы модели прохождения сигналов (например, синусоидальная или периодическая модели) дают повышенную производительность в невыровненной
 30 области (по сравнению с выровненной областью).

Следует отметить, что в альтернативном примере выход блока 517 предсказания (то есть блок 150 предполагаемых коэффициентов преобразования) может добавляться на выходе узла 114 обратного выравнивания (то есть к блоку 149 восстановленных
 35 коэффициентов) (см. фиг.23а). Тогда узел 544 формирователя из фиг.23с можно сконфигурировать для выполнения объединенной операции из задержанного выравнивания и обратного выравнивания.

Элементы в принятом потоке битов могут управлять случающейся время от времени очисткой буфера 541 субполосы и буфера 541 огибающих, например, в случае первой
 40 единицы кодирования (то есть первого блока) I-кадра. Это дает возможность декодировать I-кадр без сведений о предыдущих данных. Первая единица кодирования обычно не может использовать содействие блока предсказания, но может, тем не менее, использовать сравнительно меньшее количество битов для передачи информации 520 о блоке предсказания. Потерю усиления предсказания можно компенсировать распределением большего количества битов кодированию ошибки предсказания у этой
 45 первой единицы кодирования. Обычно содействие блока предсказания важно для второй единицы кодирования (то есть второго блока) I-кадра. Благодаря этим аспектам качество можно поддерживать при сравнительно небольшом увеличении скорости передачи битов, даже при очень частом использовании I-кадров.

Другими словами, наборы 132, 332 блоков (также называемые кадрами) содержат множество блоков 131, которое может кодироваться с использованием кодирования с предсказанием. При кодировании I-кадра только первый блок 203 из набора 332 блоков нельзя кодировать с использованием эффективности кодирования, достигаемой кодером с предсказанием. Уже непосредственно следующий блок 201 может использовать выгоды кодирования с предсказанием. Это означает, что недостатки I-кадра в отношении эффективности кодирования ограничиваются кодированием первого блока 203 коэффициентов преобразования в кадре 332 и не применяются к другим блокам 201, 204, 205 в кадре 332. Поэтому схема кодирования речи с преобразованием, описанная в настоящем документе, допускает относительно частое использование I-кадров без значительного влияния на эффективность кодирования. По существу, описываемая в настоящее время схема кодирования речи с преобразованием особенно подходит для применений, которые требуют довольно быстрой и/или довольно частой синхронизации между декодером и кодером.

Фиг.23d показывает блок-схему примерного декодера 502 спектра. Декодер 502 спектра содержит декодер 551 без потерь, который конфигурируется для декодирования энтропийно кодированных данных 163 о коэффициентах. Кроме того, декодер 502 спектра содержит обратный квантователь 552, который конфигурируется для назначения значений коэффициентов индексам квантования, содержащимся в данных 163 о коэффициентах. Как указано применительно к кодеру 100, 170, разные коэффициенты преобразования можно квантовать с использованием разных квантователей, выбранных из набора заранее установленных квантователей, например, из конечного набора скалярных квантователей, основанных на модели. Как показано на фиг.22, набор квантователей 321, 322, 323 может содержать разные типы квантователей. Набор квантователей может содержать квантователь 321, который обеспечивает синтез шума (в случае нулевой скорости передачи битов), один или несколько квантователей 322 с добавлением псевдослучайного шума (для сравнительно низких отношений сигнал-шум, SNR, и для промежуточных скоростей передачи битов) и/или один или несколько простых квантователей 323 (для сравнительно высоких SNR и для сравнительно высоких скоростей передачи битов).

Узел 107 уточнения огибающей может конфигурироваться для предоставления огибающей 138 распределения, которую можно объединить с параметром смещения, содержащимся в данных 163 о коэффициентах, чтобы получить вектор распределения. Вектор распределения содержит целое значение для каждой полосы 302 частот. Целое значение для конкретной полосы 302 частот указывает на точку искажения в зависимости от скорости передачи, используемую для обратного квантования коэффициентов преобразования в конкретной полосе 302. Другими словами, целое значение для конкретной полосы 302 частот указывает на квантователь, используемый для обратного квантования коэффициентов преобразования в конкретной полосе 302. Увеличение целого значения на единицу соответствует увеличению SNR на 1,5 дБ. Для квантователей 322 с добавлением псевдослучайного шума и простых квантователей 323 при кодировании без потерь можно использовать модель распределения вероятностей Лапласа, которая может применять арифметическое кодирование. Один или несколько квантователей 322 с добавлением псевдослучайного шума можно использовать для плавного сокращения разрыва между случаями с низкой и высокой скоростью передачи битов. Квантователи 322 с добавлением псевдослучайного шума могут быть полезны при создании достаточно ровного качества выходного аудио для стационарных шумоподобных сигналов.

Другими словами, обратный квантователь 552 может конфигурироваться для приема индексов квантования коэффициентов в текущем блоке 131 коэффициентов преобразования. Один или несколько индексов квантования коэффициентов в конкретной полосе 302 частот определены с использованием соответствующего квантователя из заранее установленного набора квантователей. Значение вектора распределения (которое может определяться путем смещения огибающей 138 распределения с помощью параметра смещения) для конкретной полосы 302 частот указывает квантователь, который использован для определения одного или нескольких индексов квантования коэффициентов в конкретной полосе 302 частот. После идентификации квантователя один или несколько индексов квантования коэффициентов можно обратно квантовать, чтобы получить блок 145 квантованных коэффициентов ошибки.

Кроме того, спектральный декодер 502 может содержать узел 113 обратного изменения масштаба, чтобы предоставить блок 147 масштабированных квантованных коэффициентов ошибки. Дополнительные инструменты и взаимосвязи вокруг декодера 551 без потерь и обратного квантователя 552 из фиг.23d могут использоваться для приспособления спектрального декодирования к использованию во всем декодере 500, показанном на фиг.23а, где выход спектрального декодера 502 (то есть блок 145 квантованных коэффициентов ошибки) используется для обеспечения аддитивной поправки в предсказанный вектор выровненной области (то есть в блок 150 предполагаемых коэффициентов преобразования). В частности, дополнительные инструменты могут обеспечивать, что выполняемая декодером 500 обработка соответствует обработке, выполняемой кодером 100, 170.

В частности, спектральный декодер 502 может содержать узел 111 эвристического масштабирования. Как показано в сочетании с кодером 100, 170, узел 111 эвристического масштабирования может влиять на распределение битов. В кодере 100, 170 текущие блоки 141 коэффициентов ошибки предсказания можно масштабировать вплоть до единичной дисперсии с помощью эвристического правила. В результате распределение по умолчанию может привести к слишком мелкому квантованию окончательного результата с уменьшенным масштабом из узла 111 эвристического масштабирования. Поэтому распределение следует изменить аналогично изменению коэффициентов ошибки предсказания.

Однако, как указано ниже, может быть полезно избежать сокращения ресурсов кодирования для одного или нескольких элементов разрешения низкой частоты (или полос низких частот). В частности, это может быть полезно для борьбы с артефактом (низкочастотного) рокота/шума, который оказывается наиболее заметным в вокализованных ситуациях (то есть для сигнала, имеющего сравнительно большой управляющий параметр 146, g_{fu}). По существу, распределение битов/выбор квантователя в зависимости от управляющего параметра 146, который описывается ниже, может считаться "адаптивным к речи подъемом качества на низких частотах".

Спектральный декодер может зависеть от управляющего параметра 146, называемого g_{fu} , который является ограниченной версией усиления g блока предсказания, $g_{fu} = \min(1, (\max(g, 0)))$.

С использованием управляющего параметра 146 можно адаптировать набор квантователей, используемый в узле 112 квантования коэффициентов в кодере 100, 170 и используемый в обратном квантователе 552. В частности, на основе управляющего параметра 146 можно адаптировать зашумленность набора квантователей. В качестве примера близкое к 1 значение управляющего параметра 146, g_{fu} , может инициировать

ограничение диапазона уровней распределения, использующих квантователи с добавлением псевдослучайного шума, и может инициировать уменьшение дисперсии уровня синтеза шума. В примере можно установить порог решения по псевдослучайному шуму при $rfu = 0,75$ и усилению шума, равному $1-rfu$. Адаптация псевдослучайного шума может влиять на декодирование без потерь и обратный квантователь, тогда как адаптация усиления шума обычно влияет только на обратный квантователь.

Можно предположить, что содействие блока предсказания важно для вокализованных/тональных ситуаций. По существу, сравнительно большое усиление g блока предсказания (то есть сравнительно большой управляющий параметр 146) может указывать вокализованный или тональный речевой сигнал. В таких ситуациях опытным путем выявлено, что добавление связанного с псевдослучайным шумом или явного (случай нулевого распределения) шума непродуктивно в отношении воспринимаемого качества кодированного сигнала. В результате количество квантователей 322 с добавлением псевдослучайного шума и/или тип шума, используемый для квантователя 321 с синтезом шума, можно адаптировать на основе усиления g блока предсказания, посредством этого повышая воспринимаемое качество кодированного речевого сигнала.

По существу, управляющий параметр 146 может использоваться для изменения диапазона 324, 325 SNR, для которого используются квантователи 322 с добавлением псевдослучайного шума. В качестве примера, если управляющий параметр 146 $rfu < 0,75$, то можно использовать диапазон 324 для квантователей с добавлением псевдослучайного шума. Другими словами, если управляющий параметр 146 меньше заранее установленной пороговой величины, то можно использовать первый набор 326 квантователей. С другой стороны, если управляющий параметр 146 $rfu \geq 0,75$, то можно использовать диапазон 325 для квантователей с добавлением псевдослучайного шума. Другими словами, если управляющий параметр 146 больше либо равен заранее установленной пороговой величине, то можно использовать второй набор 327 квантователей.

Кроме того, управляющий параметр 146 может использоваться для изменения дисперсии и распределения битов. Причина в том, что успешное предсказание обычно потребует меньшей коррекции, особенно в низкочастотном диапазоне от 0 до 1 кГц. Может быть выгодно явно сообщить квантователю об этом отклонении от модели единичной дисперсии, чтобы освободить ресурсы кодирования для полос 302 высоких частот.

Эквиваленты, расширения, альтернативы и прочее

Дополнительные варианты осуществления настоящего изобретения станут ясны специалисту в данной области техники после изучения вышеприведенного описания. Даже если настоящее описание и чертежи раскрывают варианты осуществления и примеры, изобретение не ограничивается этими конкретными примерами. Можно вносить многочисленные модификации и изменения без отклонения от объема настоящего изобретения, который задается прилагаемой формулой изобретения. Никакие ссылочные позиции, наблюдаемые в формуле изобретения, не должны восприниматься как ограничивающие ее объем.

Раскрытые выше системы и способы можно реализовать в виде программного обеспечения, микропрограммного обеспечения, аппаратных средств или их сочетания. При аппаратной реализации разделение задач между функциональными узлами, упоминаемыми в вышеприведенном описании, не обязательно соответствует разделению на физические узлы; наоборот, один физический компонент может обладать несколькими

функциональными возможностями, и одна задача может осуществляться несколькими физическими компонентами совместно. Некоторые компоненты или все компоненты можно реализовать в виде программного обеспечения, исполняемого цифровым процессором сигналов или микропроцессором, либо можно реализовать в виде аппаратных средств или в виде специализированной интегральной схемы. Такое программное обеспечение может распространяться на машиночитаемых носителях, которые могут быть выполнены в виде компьютерных носителей информации (или постоянных носителей) и средств связи (или временных носителей). Как известно специалисту в данной области техники, термин "компьютерные носители информации" включает в себя энергозависимые и энергонезависимые, съемные и несъемные носители, реализованные по любому способу или технологии для хранения информации, например машиночитаемых команд, структур данных, программных модулей или других данных. Компьютерные носители информации включают в себя, но не ограничиваются, RAM, ROM, EEPROM, флэш-память или другую технологию памяти, компакт-диск, универсальные цифровые диски (DVD) или другой накопитель на оптических дисках, магнитные кассеты, магнитную ленту, накопитель на магнитных дисках или другие магнитные запоминающие устройства, или любой другой носитель, который может использоваться для хранения нужной информации и к которому можно обращаться с помощью компьютера. Кроме того, специалисту в данной области техники известно, что средства связи обычно воплощают машиночитаемые команды, структуры данных, программные модули или другие данные в модулированном сигнале данных, таком как несущая или другой транспортный механизм, и включают в себя любые средства доставки информации.

(57) Формула изобретения

1. Система обработки аудио (100), сконфигурированная для приема аудиопотока битов, причем система обработки аудио содержит:

декодер (101), приспособленный для приема потока битов и вывода квантованных спектральных коэффициентов;

компонент предварительной обработки, который включает в себя:

- каскад деквантования (102), приспособленный для приема квантованных спектральных коэффициентов и для вывода первого представления частотной области для промежуточного сигнала; и

- каскад обратного преобразования (103) для приема первого представления частотной области для промежуточного сигнала и синтеза на его основе представления временной области для промежуточного сигнала;

каскад обработки, который включает в себя:

- гребенку фильтров анализа (104) для приема представления временной области для промежуточного сигнала и вывода второго представления частотной области для промежуточного сигнала;

- по меньшей мере один компонент обработки (105, 106, 107) для приема упомянутого второго представления частотной области для промежуточного сигнала и вывода представления частотной области для обработанного аудиосигнала; и

- гребенку фильтров синтеза (108) для приема представления частотной области для обработанного аудиосигнала и вывода представления временной области для обработанного аудиосигнала; и

преобразователь частоты дискретизации (109) для приема упомянутого представления временной области для обработанного аудиосигнала и вывода восстановленного

аудиосигнала, дискретизированного с целевой частотой дискретизации,

причем соответствующие внутренние частоты дискретизации представления временной области для промежуточного аудиосигнала и представления временной области для обработанного аудиосигнала равны, и причем упомянутый по меньшей мере один компонент обработки включает в себя:

каскад параметрического повышающего микширования (106) для приема сигнала понижающего микширования с M каналами и вывода на его основе сигнала с N каналами, причем каскад параметрического повышающего микширования работает по меньшей мере в режиме, где $1 \leq M < N$, ассоциированном с задержкой, и в режиме, где $1 \leq M = N$; и

первый каскад задержки, сконфигурированный для вызова задержки, когда каскад параметрического повышающего микширования находится в режиме, где $1 \leq M = N$, чтобы компенсировать задержку, ассоциированную с режимом, где $1 \leq M < N$, чтобы каскад обработки имел постоянную суммарную задержку независимо от текущего режима работы каскада параметрического повышающего микширования.

2. Система обработки аудио по п.1, в которой компонент предварительной обработки работает в аудиорежиме и характерном для речи режиме и в которой изменение режима с аудиорежима на характерный для речи режим у компонента предварительной обработки включает в себя сокращение максимальной длины кадра у каскада обратного преобразования.

3. Система обработки аудио по п.2, в которой преобразователь частоты дискретизации выполнен с возможностью предоставления восстановленного аудиосигнала, дискретизированного с целевой частотой дискретизации, отличающейся до 5% от внутренней частоты дискретизации упомянутого представления временной области для обработанного аудиосигнала.

4. Система обработки аудио по п.1, дополнительно содержащая обходную линию, размещенную параллельно каскаду обработки и содержащую второй каскад задержки, сконфигурированный для вызова задержки, равной постоянной суммарной задержке у каскада обработки.

5. Система обработки аудио по п.1, в которой каскад параметрического повышающего микширования дополнительно работает по меньшей мере в режиме, где $M = 3$ и $N = 5$.

6. Система обработки аудио по п.5, в которой компонент предварительной обработки конфигурируется для предоставления промежуточного сигнала, содержащего сигнал понижающего микширования, в том режиме каскада параметрического повышающего микширования, где $M = 3$ и $N = 5$, причем компонент предварительной обработки выводит два канала среди $M = 3$ каналов из совместно кодированных каналов в аудиопотоке битов.

7. Система обработки аудио по п.1, в которой упомянутый по меньшей мере один компонент обработки дополнительно включает в себя модуль копирования спектральных полос (106), размещенный раньше каскада параметрического повышающего микширования и выполненный с возможностью восстановления высокочастотного содержимого, причем модуль копирования спектральных полос

- сконфигурирован быть активным по меньшей мере в тех режимах каскада

параметрического повышающего микширования, где $M < N$; и

- работает независимо от текущего режима каскада параметрического повышающего микширования, когда каскад параметрического повышающего микширования находится в любом из режимов, где $M = N$.

8. Система обработки аудио по п.7, в которой упомянутый по меньшей мере один компонент обработки дополнительно включает в себя каскад кодирования по форме, размещенный параллельно или после каскада параметрического повышающего микширования и работающий для дополнения каждого из N каналов кодированным по форме низкочастотным содержимым, причем каскад кодирования по форме включается и отключается независимо от текущего режима каскада параметрического повышающего микширования и модуля копирования спектральных полос.

9. Система обработки аудио по п.8, работающая по меньшей мере в режиме декодирования, причем каскад параметрического повышающего микширования находится в режиме $M = N$ при $M > 2$.

10. Система обработки аудио по п.9, работающая по меньшей мере в следующих режимах декодирования:

i) каскад параметрического повышающего микширования в режиме $M = N = 1$;

ii) каскад параметрического повышающего микширования в режиме $M = N = 1$, и модуль копирования спектральных полос активен;

iii) каскад параметрического повышающего микширования в режиме $M = 1, N = 2$, и модуль копирования спектральных полос активен;

iv) каскад параметрического повышающего микширования в режиме $M = 1, N = 2$, модуль копирования спектральных полос активен и каскад кодирования по форме активен;

v) каскад параметрического повышающего микширования в режиме $M = 2, N = 5$, и модуль копирования спектральных полос активен;

vi) каскад параметрического повышающего микширования в режиме $M = 2, N = 5$, модуль копирования спектральных полос активен и каскад кодирования по форме активен;

vii) каскад параметрического повышающего микширования в режиме $M = 3, N = 5$, и модуль копирования спектральных полос активен;

viii) каскад параметрического повышающего микширования в режиме $M = N = 2$;

ix) каскад параметрического повышающего микширования в режиме $M = N = 2$, и модуль копирования спектральных полос активен;

x) каскад параметрического повышающего микширования в режиме $M = N = 7$;

xi) каскад параметрического повышающего микширования в режиме $M = N = 7$, и модуль копирования спектральных полос активен.

11. Система обработки аудио по п.1, дополнительно содержащая следующие компоненты, размещенные после каскада обработки:

компонент фазового сдвига, сконфигурированный для приема представления временной области для обработанного аудиосигнала, в котором по меньшей мере один канал представляет канал окружения, и для выполнения 90-градусного фазового сдвига по меньшей мере для одного упомянутого канала окружения; и

компонент понижающего микширования, сконфигурированный для приема обработанного аудиосигнала от компонента фазового сдвига и для вывода на его основе сигнала понижающего микширования с двумя каналами.

12. Система обработки аудио по любому из предшествующих пунктов, дополнительно содержащая декодер Lfe , сконфигурированный для подготовки по меньшей мере одного дополнительного канала на основе аудиопотока битов и включения упомянутого дополнительного канала (каналов) в восстановленный аудиосигнал.

13. Способ обработки аудиопотока битов, содержащий этапы, на которых: предоставляют квантованные спектральные коэффициенты на основе потока битов;

принимают квантованные спектральные коэффициенты и выполняют обратное квантование с последующим частотно-временным преобразованием, при помощи чего получается представление временной области для промежуточного аудиосигнала;

предоставляют представление частотной области для промежуточного аудиосигнала на основе представления временной области для промежуточного аудиосигнала;

предоставляют представление частотной области для обработанного аудиосигнала путем выполнения по меньшей мере одного этапа обработки над представлением частотной области для промежуточного аудиосигнала;

предоставляют представление временной области для обработанного аудиосигнала на основе представления частотной области для обработанного аудиосигнала; и

меняют частоту дискретизации представления временной области для обработанного аудиосигнала на целевую частоту дискретизации, при помощи чего получается восстановленный аудиосигнал,

причем соответствующие внутренние частоты дискретизации представления временной области для промежуточного аудиосигнала и представления временной области для обработанного аудиосигнала равны, причем способ дополнительно содержит этап, на котором:

определяют текущий режим по меньшей мере среди режима, где $1 \leq M < N$, ассоциированного с задержкой, и режима, где $1 \leq M = N$,

причем по меньшей мере один этап обработки включает в себя:

прием сигнала понижающего микширования с M каналами и вывод на его основе сигнала с N каналами;

вызов задержки для компенсации задержки, ассоциированной с режимом, где $1 \leq M < N$, в ответ на текущий режим, являющийся режимом, где $1 \leq M = N$, чтобы суммарная задержка этапа обработки была постоянной независимо от текущего режима.

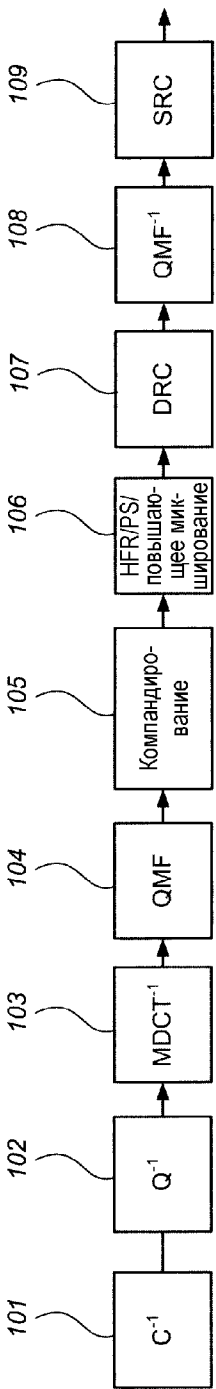
14. Способ по п.13, в котором упомянутое обратное квантование и/или частотно-временное преобразование выполняются в аппаратном компоненте, работающем по меньшей мере в аудиорежиме и характерном для речи режиме, причем текущий режим выбирается в соответствии с метаданными, ассоциированными с квантованными спектральными коэффициентами, и в котором изменение режима с аудиорежима на характерный для речи режим включает в себя сокращение максимальной длины кадра у частотно-временного преобразования.

15. Машиночитаемый носитель, содержащий компьютерную программу, сохраненную на нем, которая содержит команды для побуждения компьютера выполнять способ по любому из пп. 13-14.

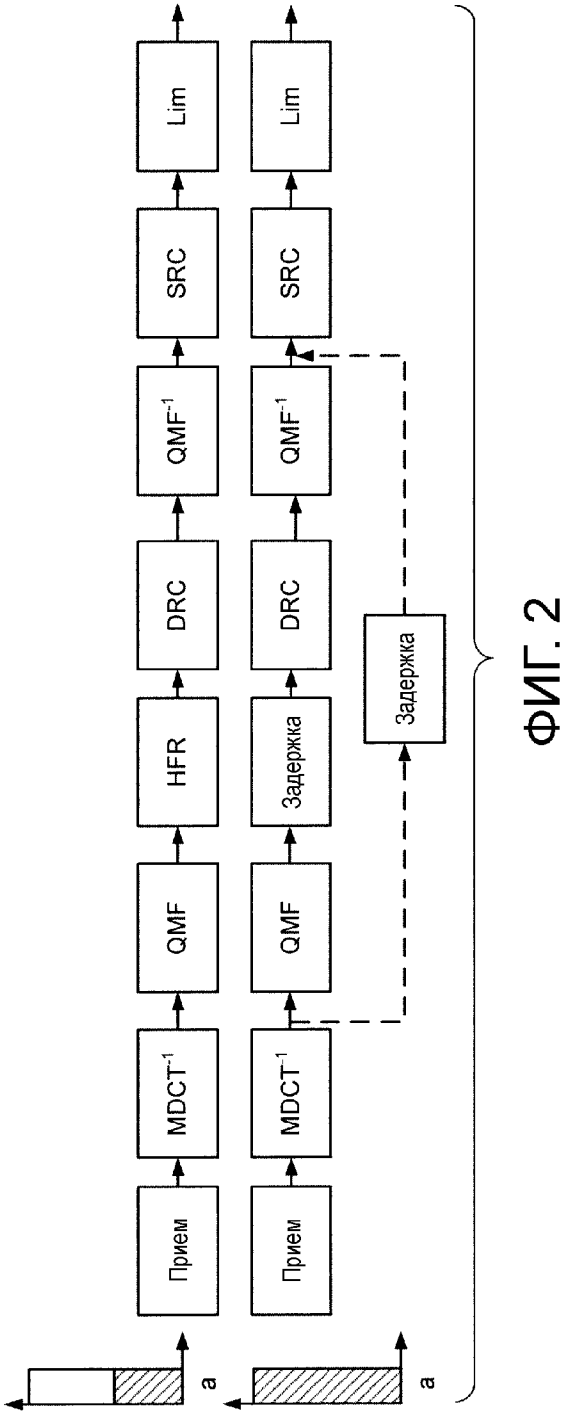
528851

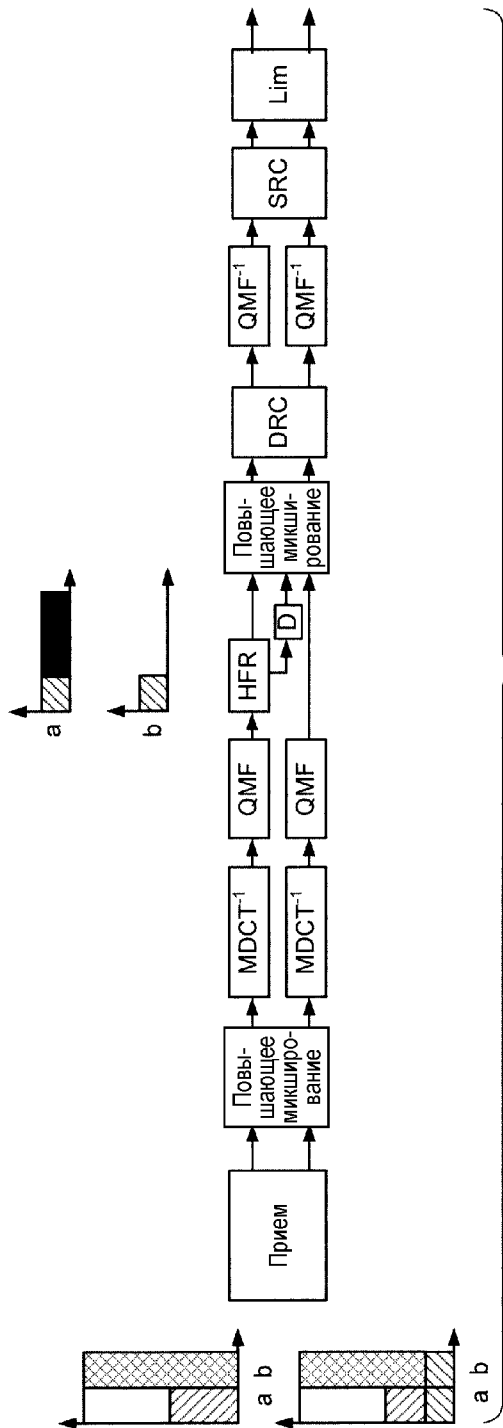
1/25

100

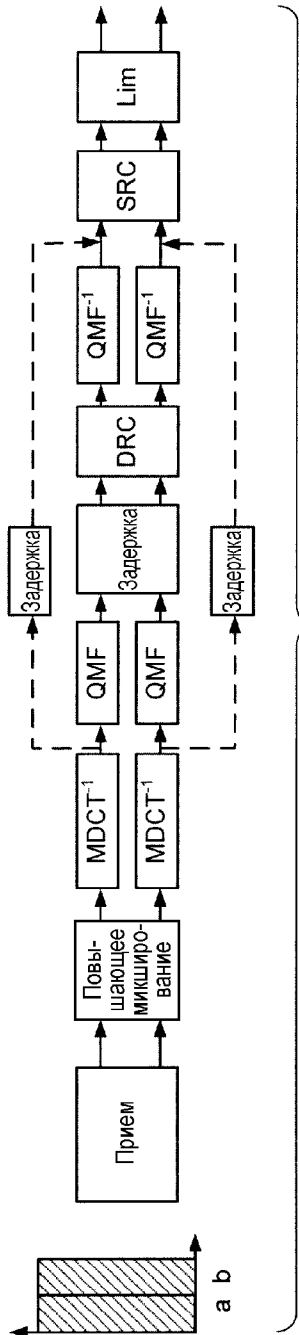


ФИГ. 1

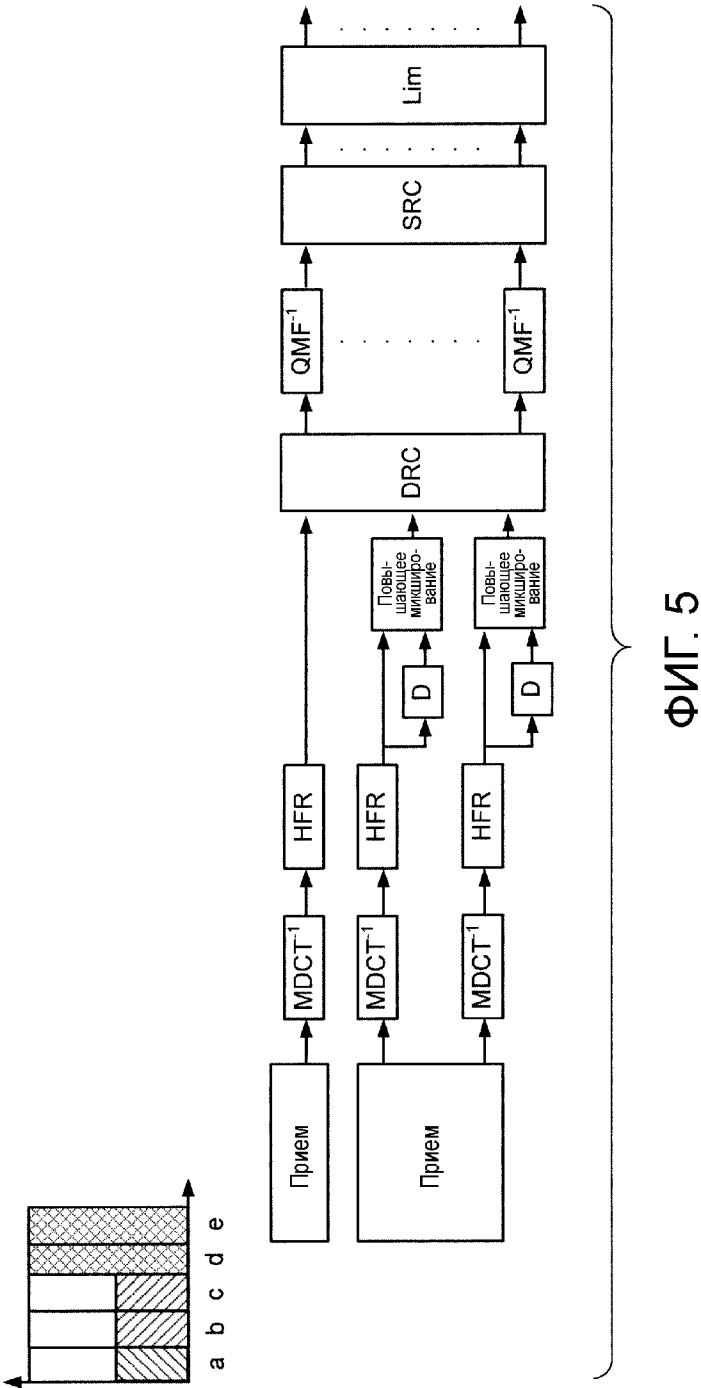




ФИГ. 3

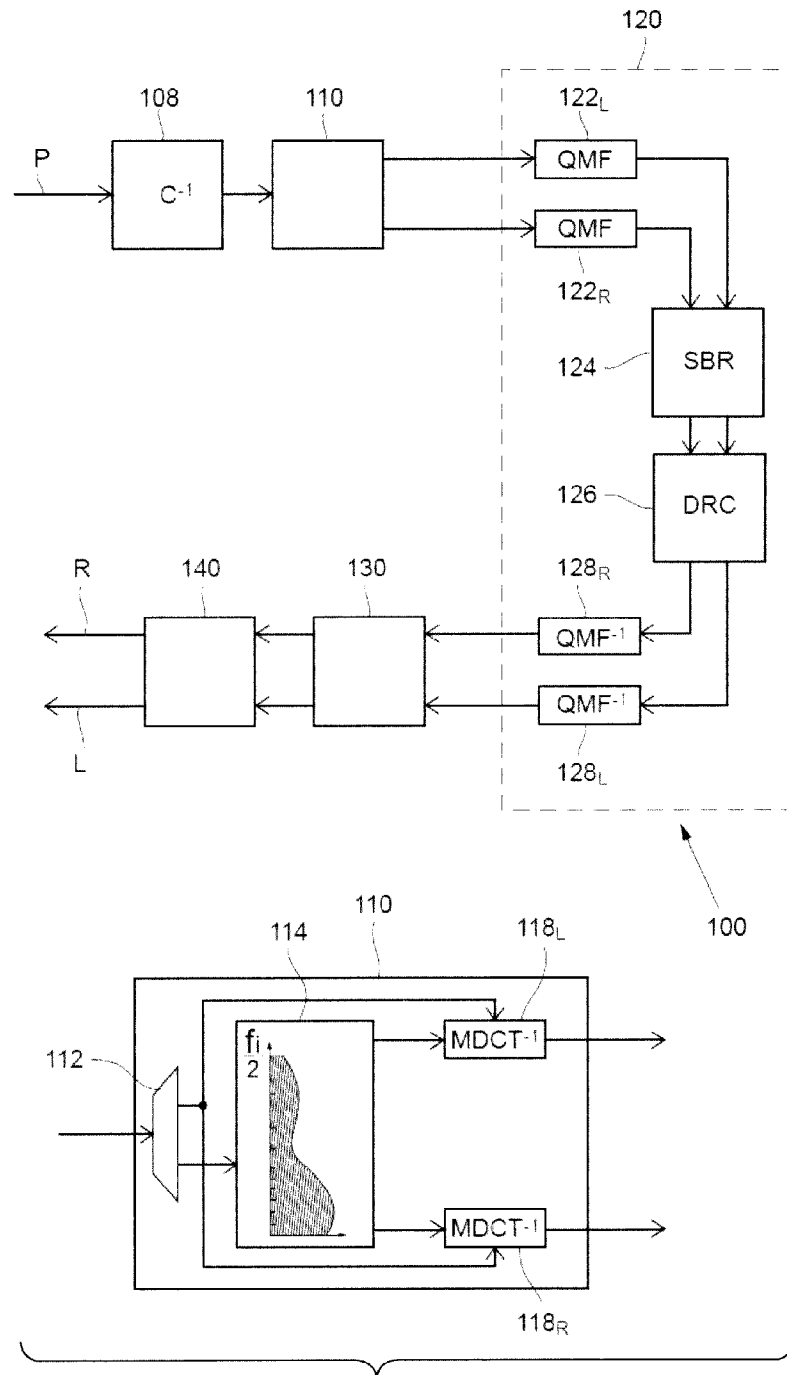


ФИГ. 4



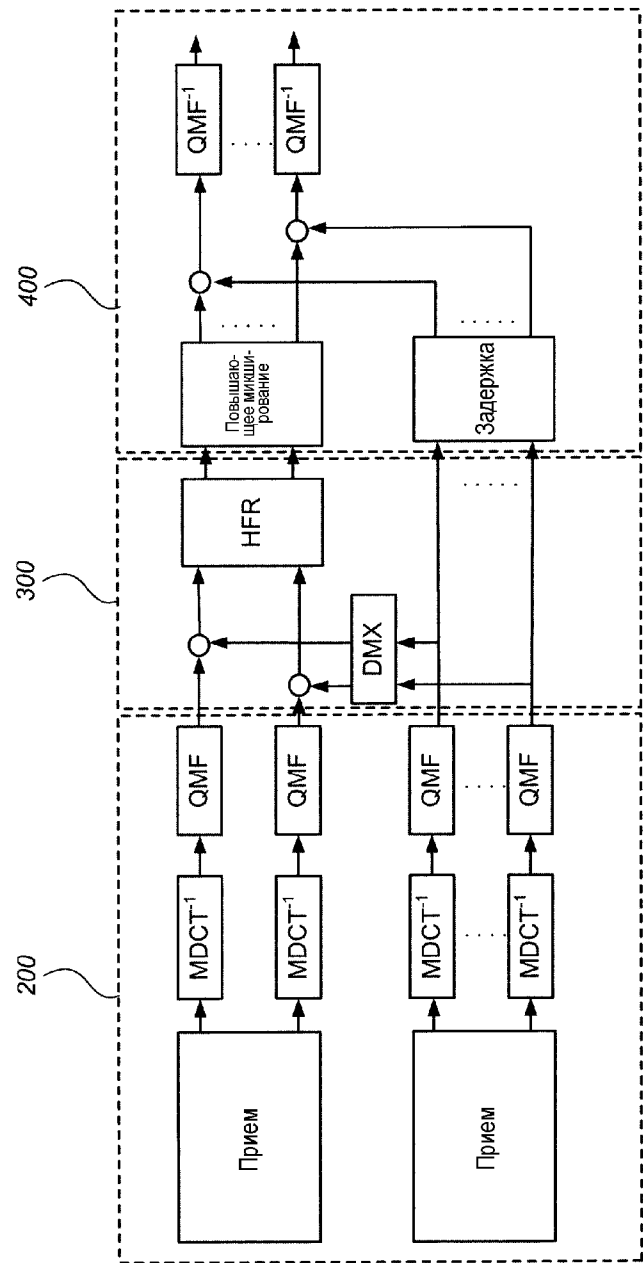
ФИГ. 5

4/25

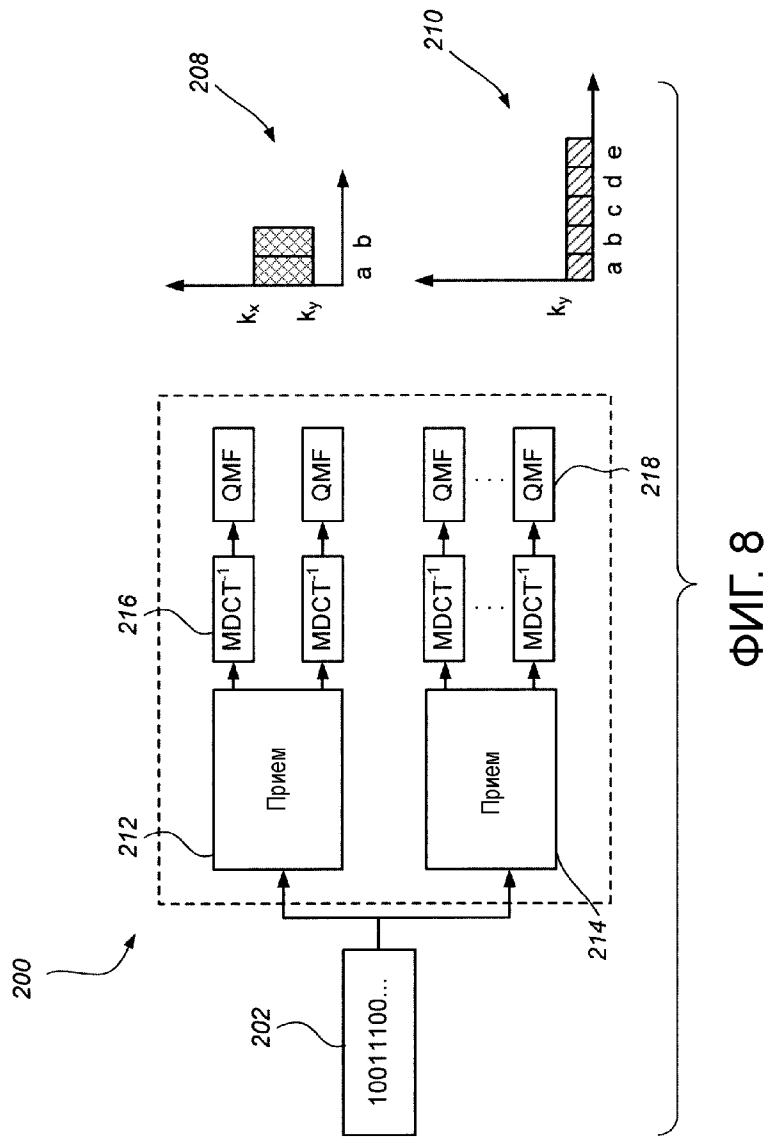


ФИГ. 6

100

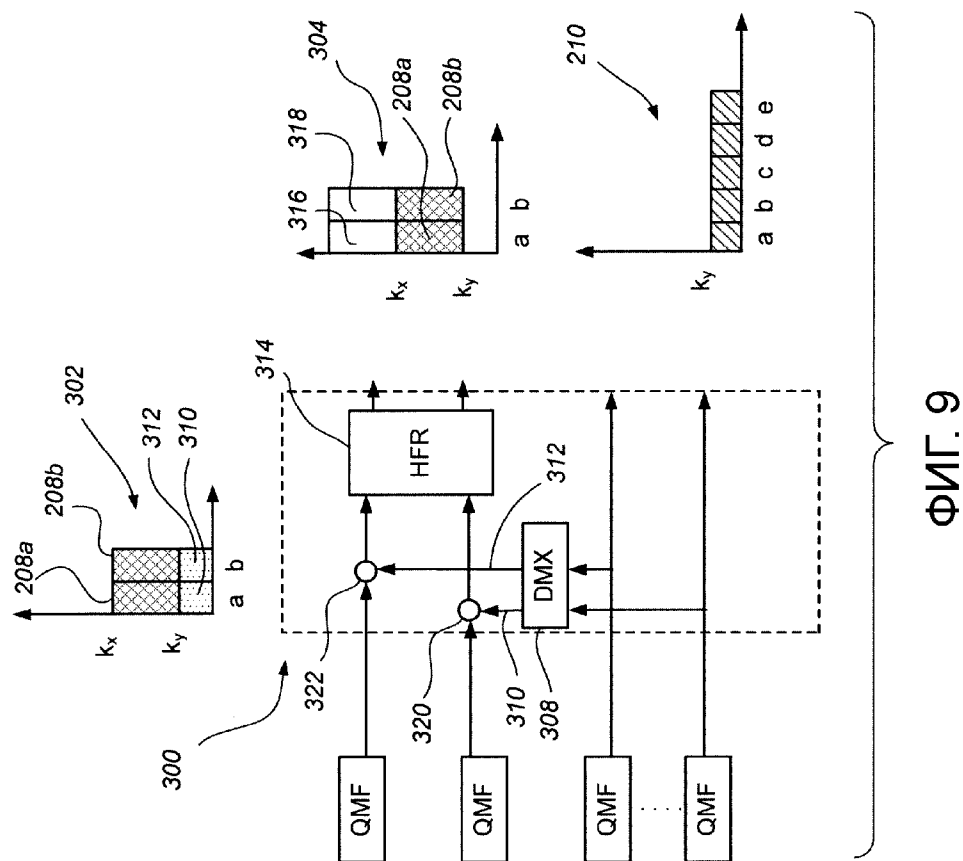


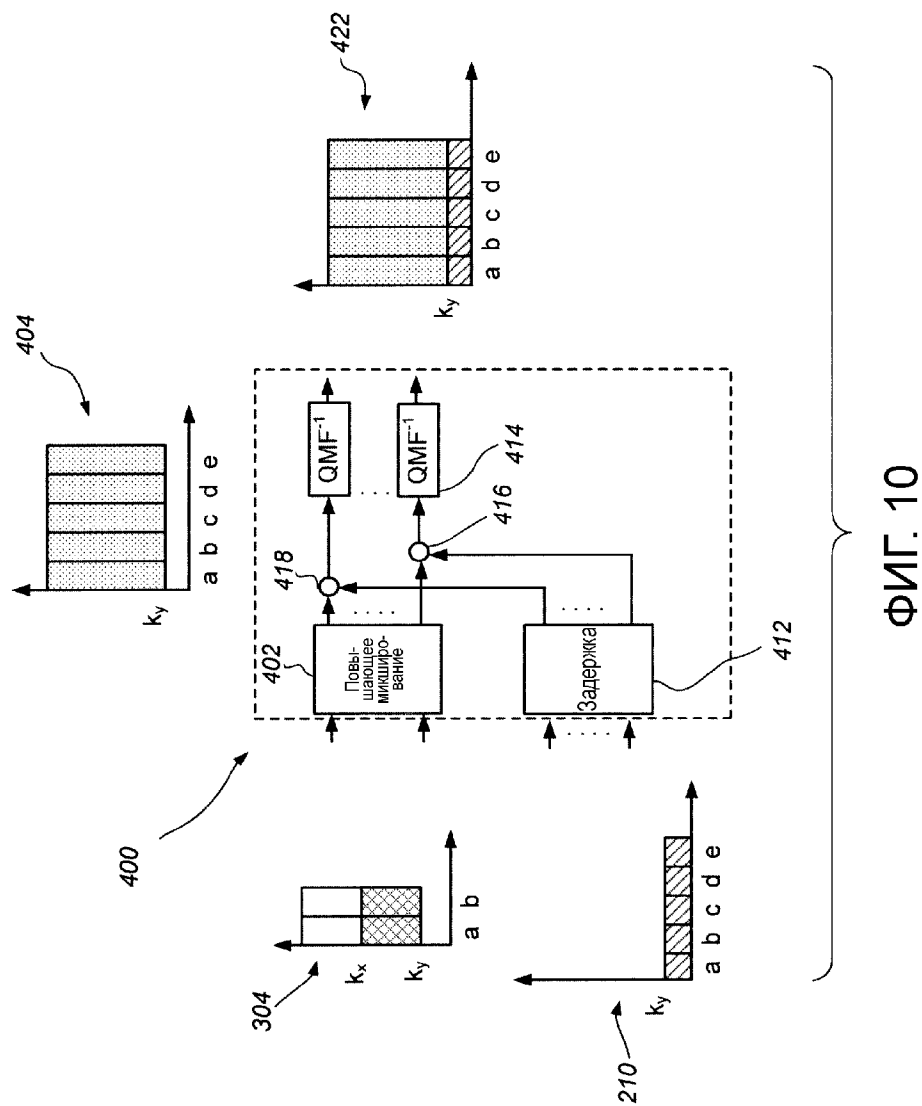
ФИГ. 7

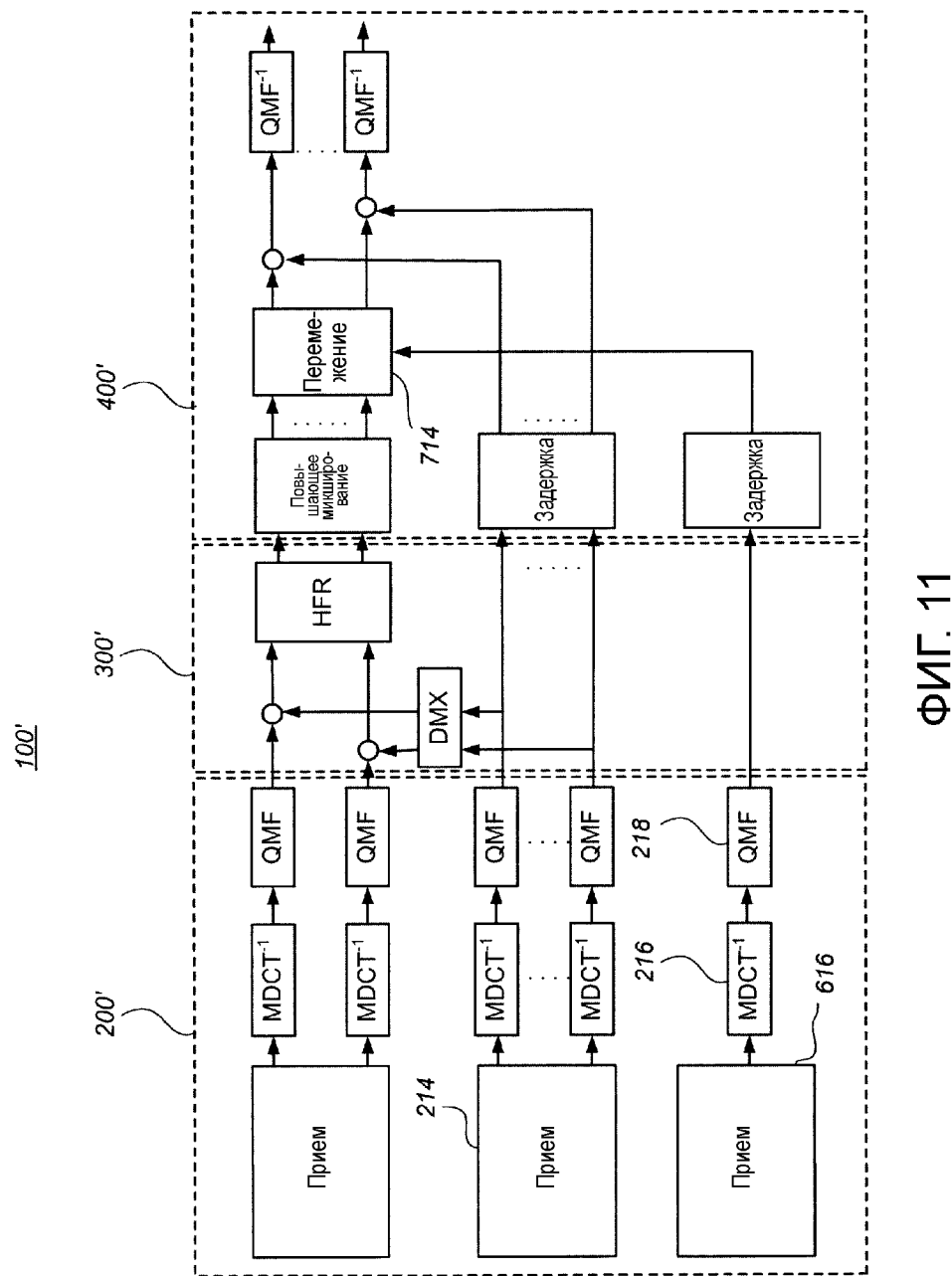


ФИГ. 8

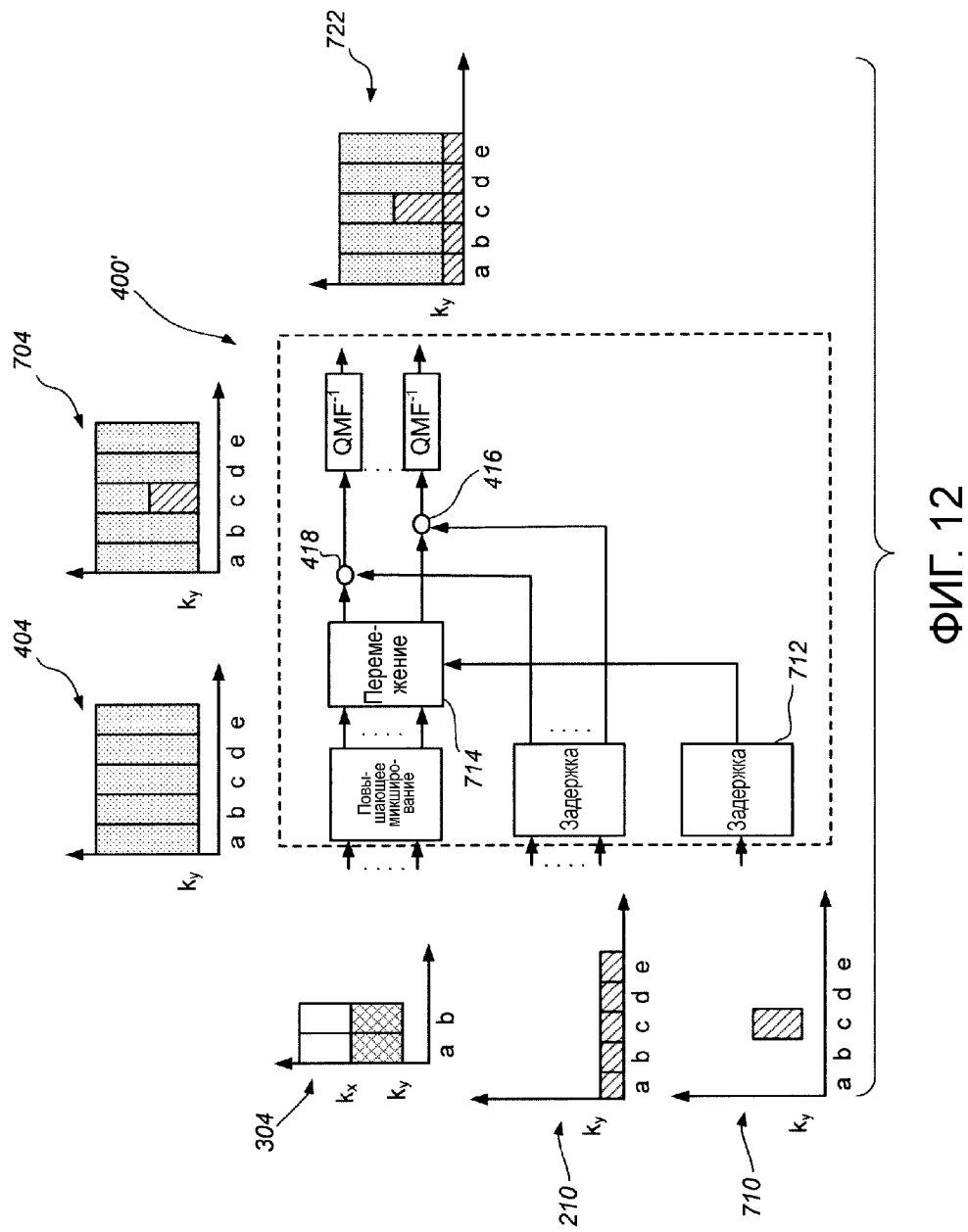
7/25



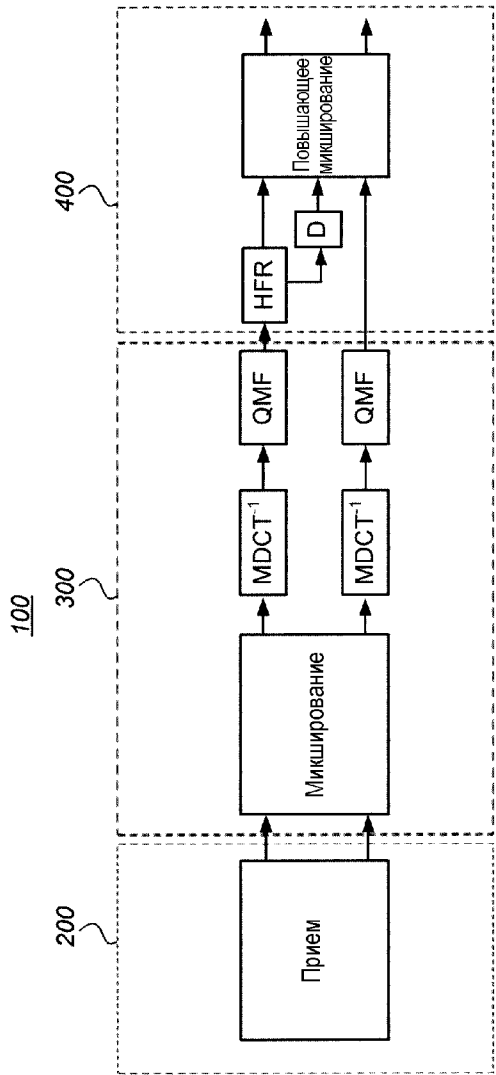




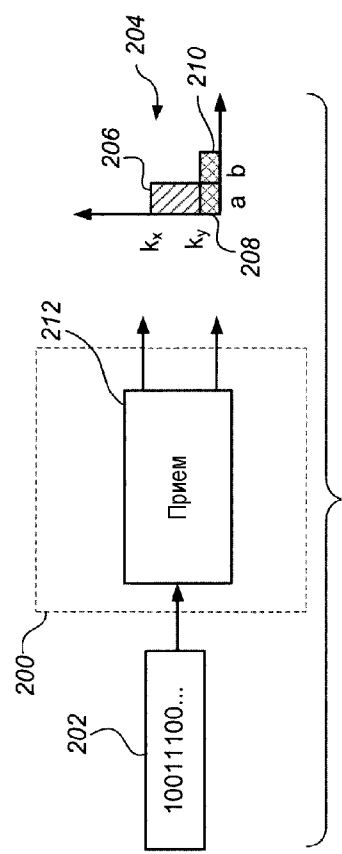
10/25



ФИГ. 12

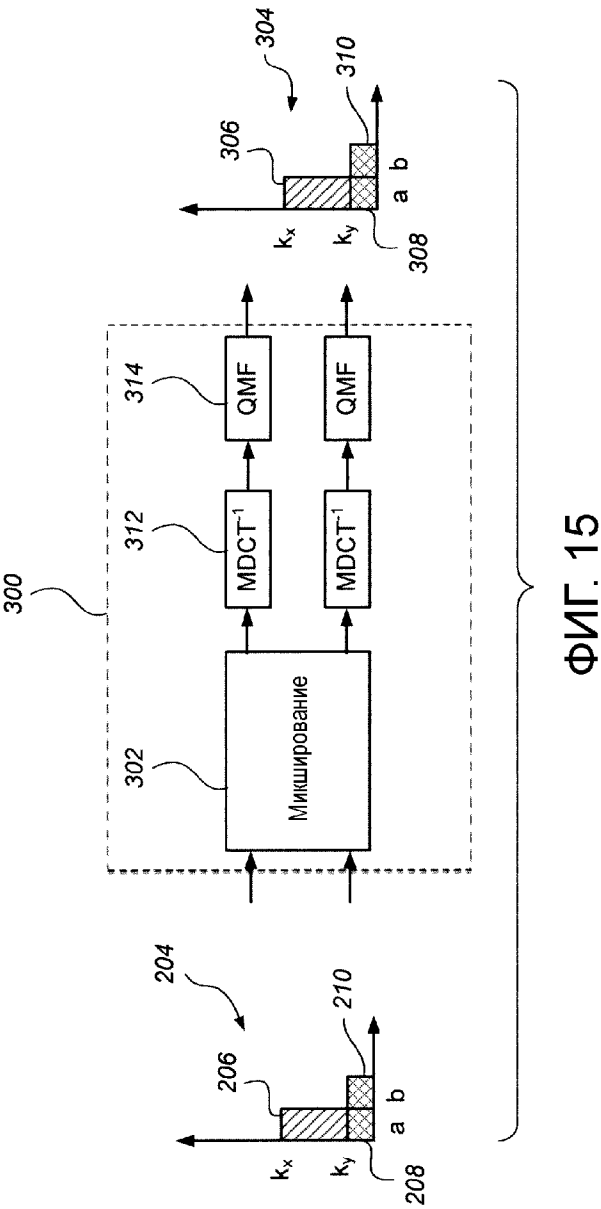


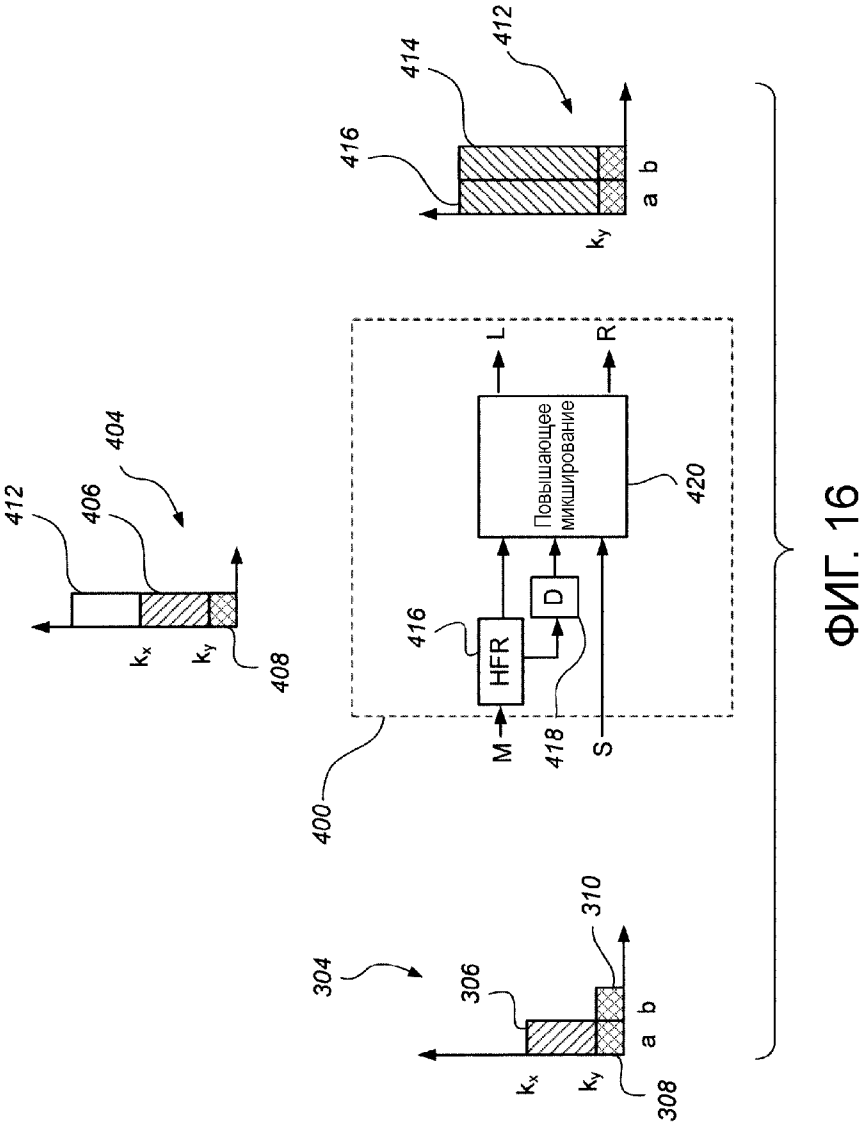
ФИГ. 13



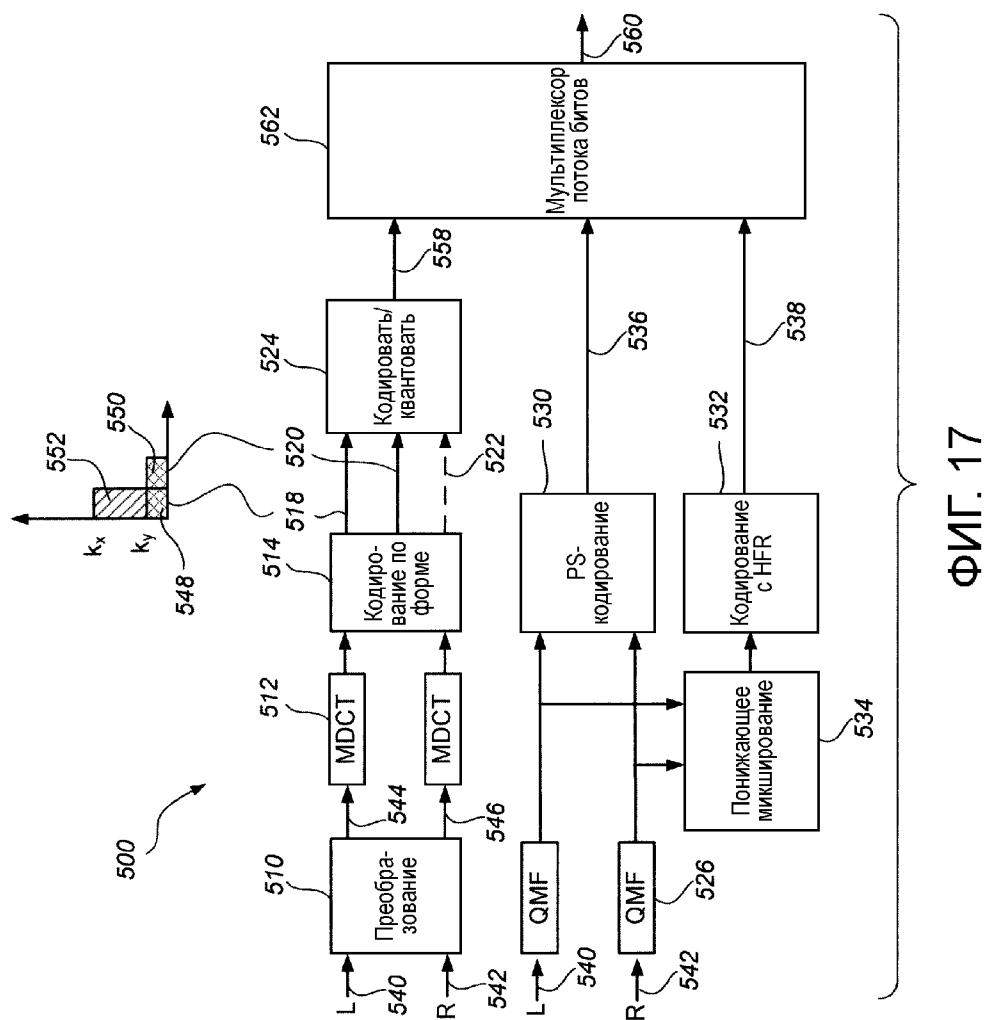
ФИГ. 14

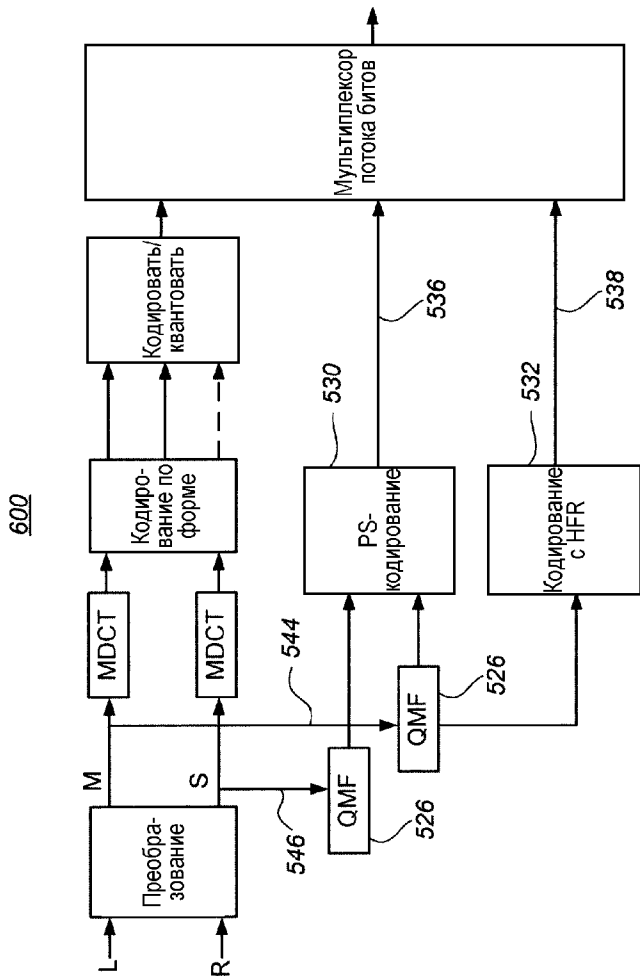
12/25



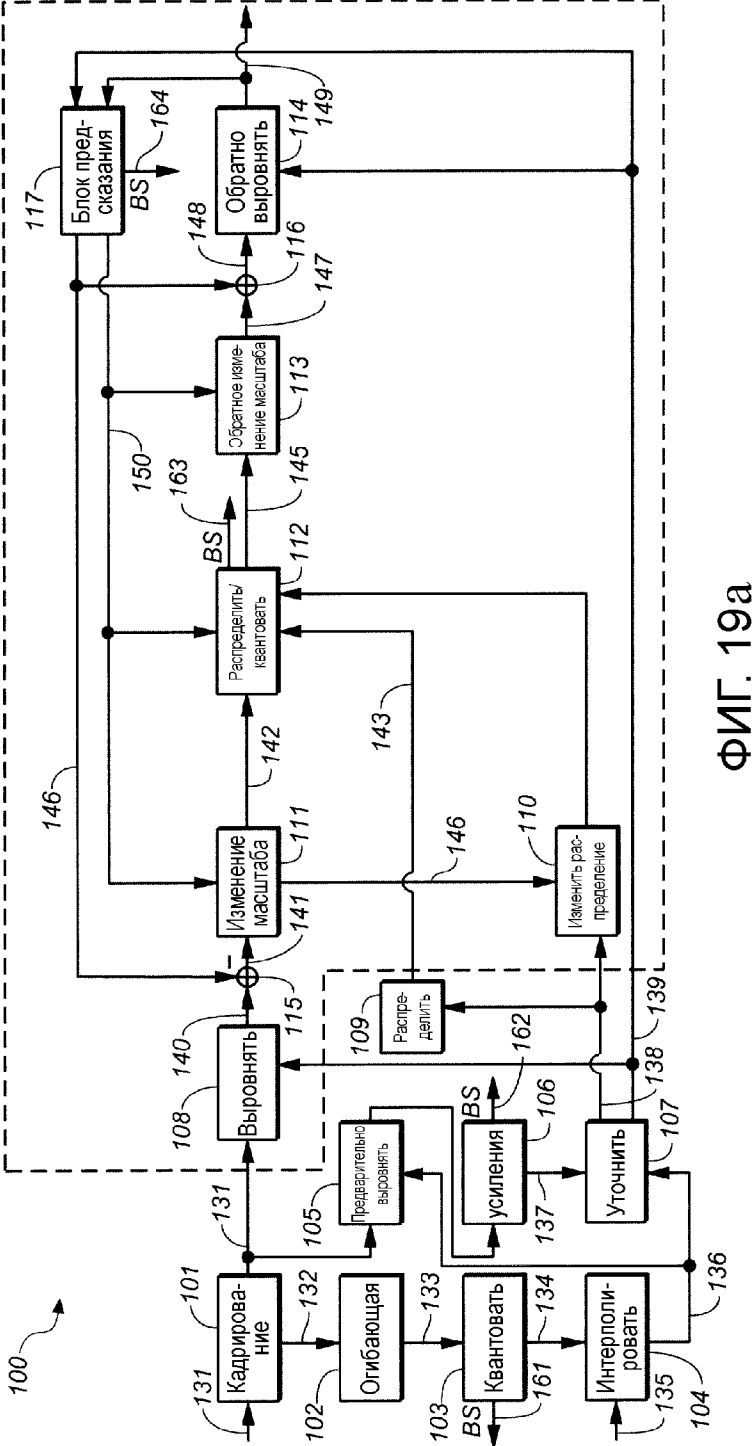


14/25



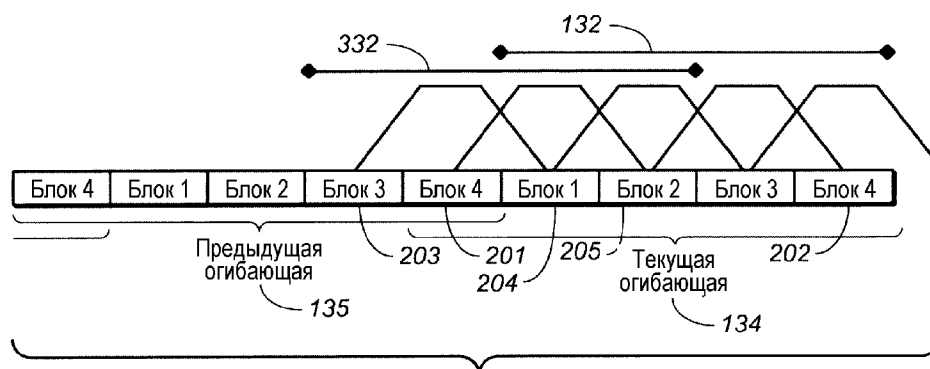


ФИГ. 18

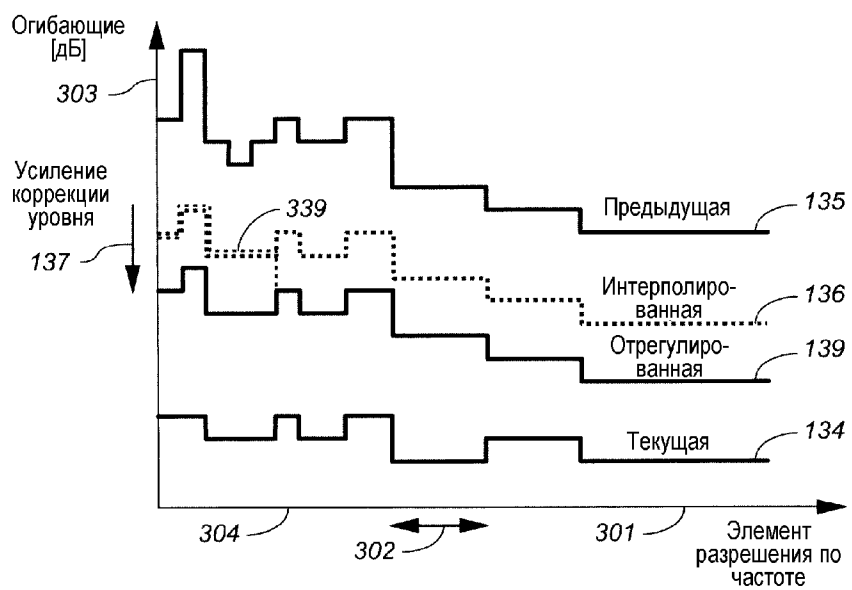




18/25

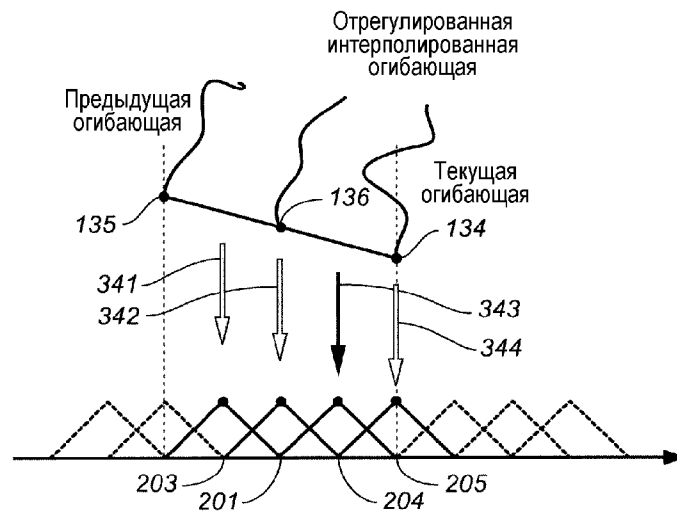


ФИГ. 20

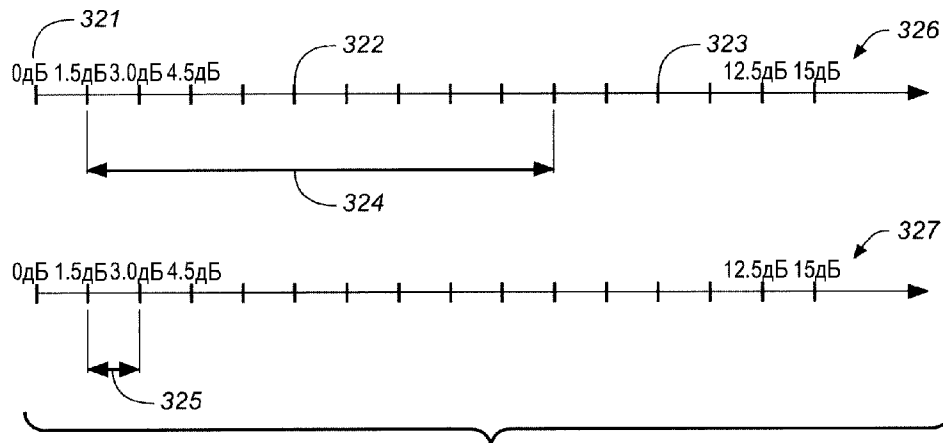


ФИГ. 21а

19/25

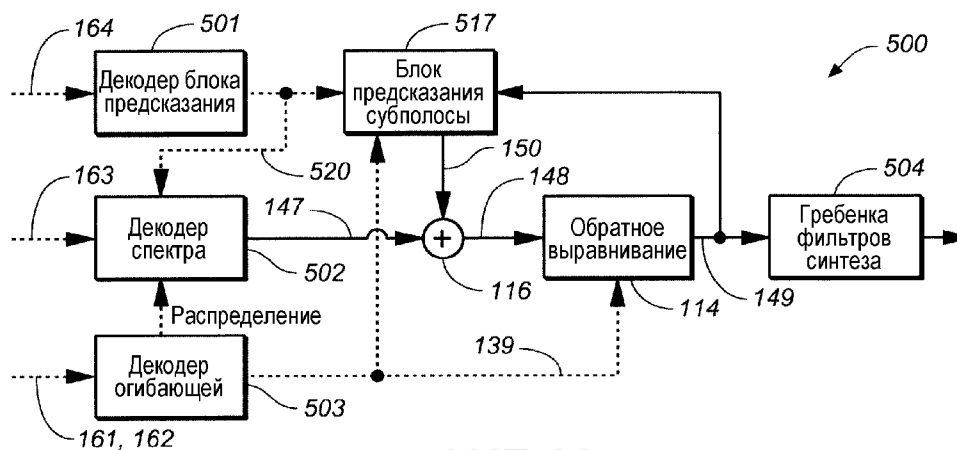


ФИГ. 21b

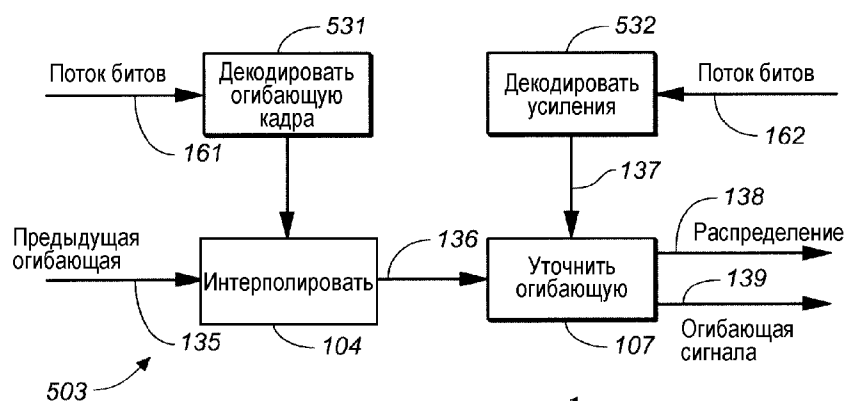


ФИГ. 22

20/25

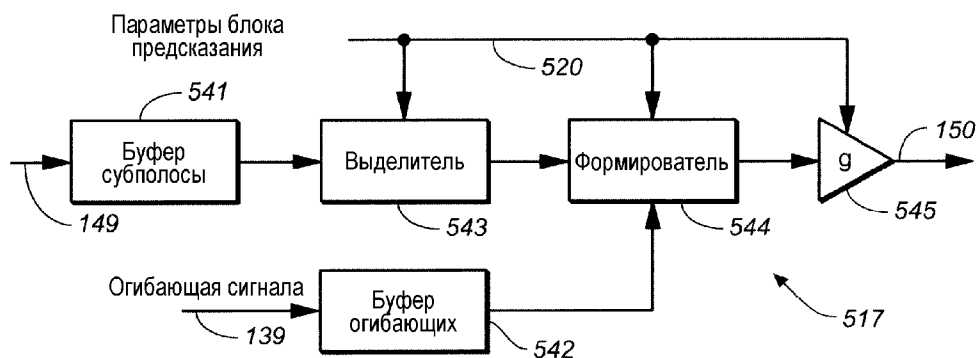


ФИГ. 23а

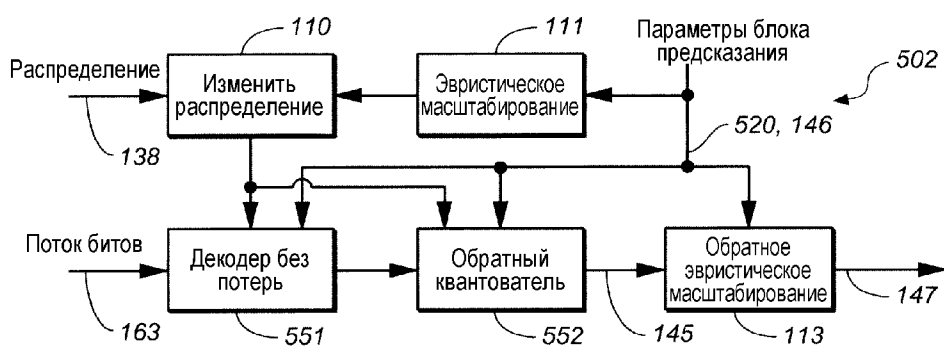


ФИГ. 23b

21/25

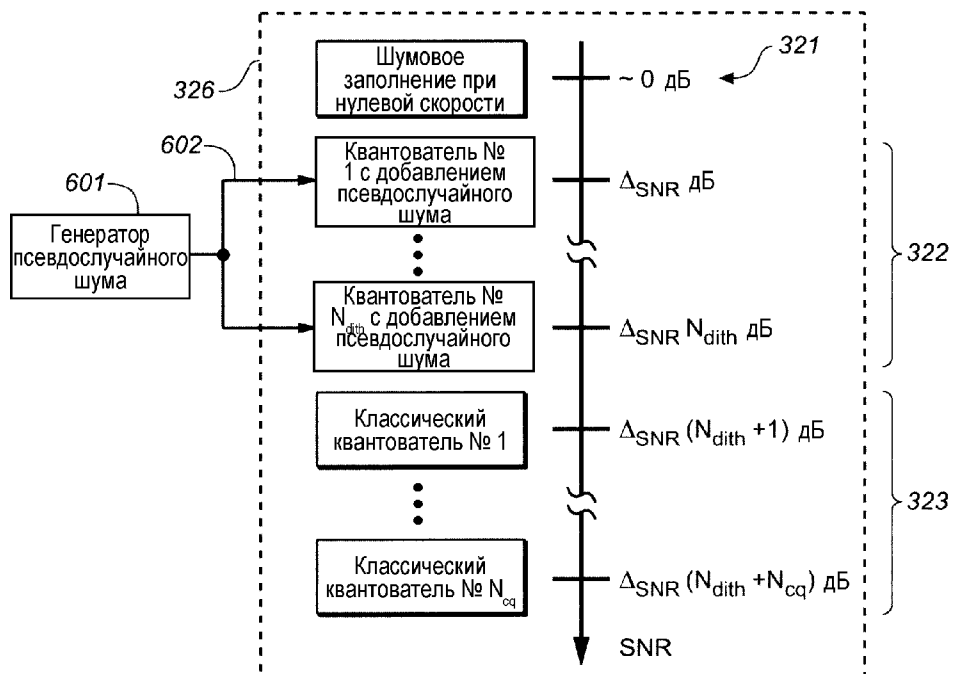


ФИГ. 23с

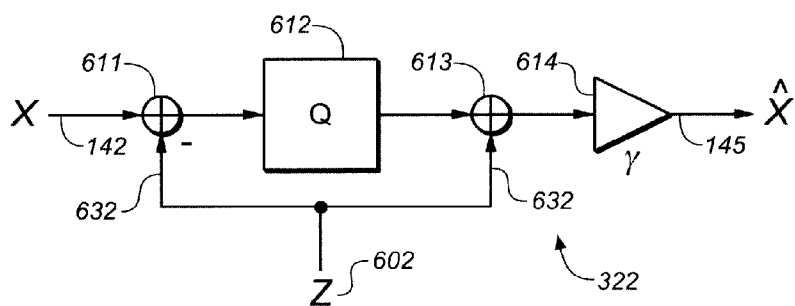


ФИГ. 23d

22/25

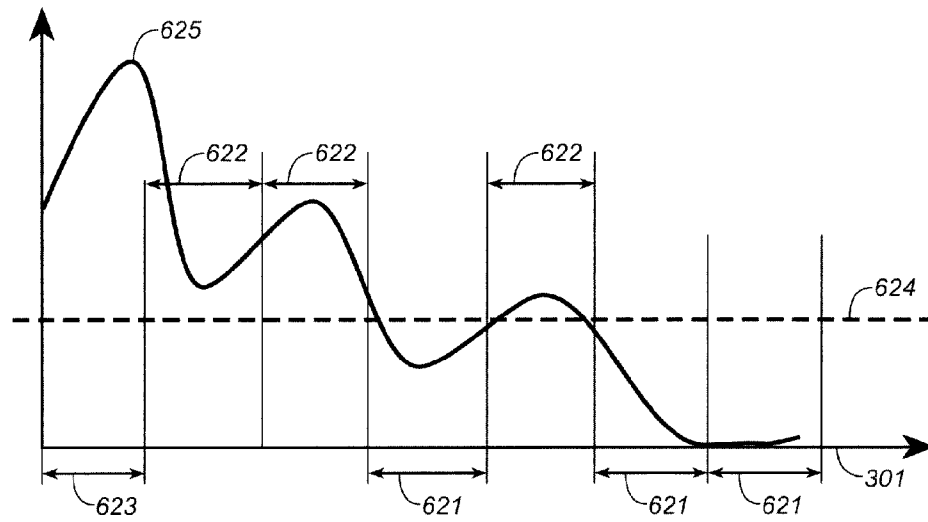


ФИГ. 24 а



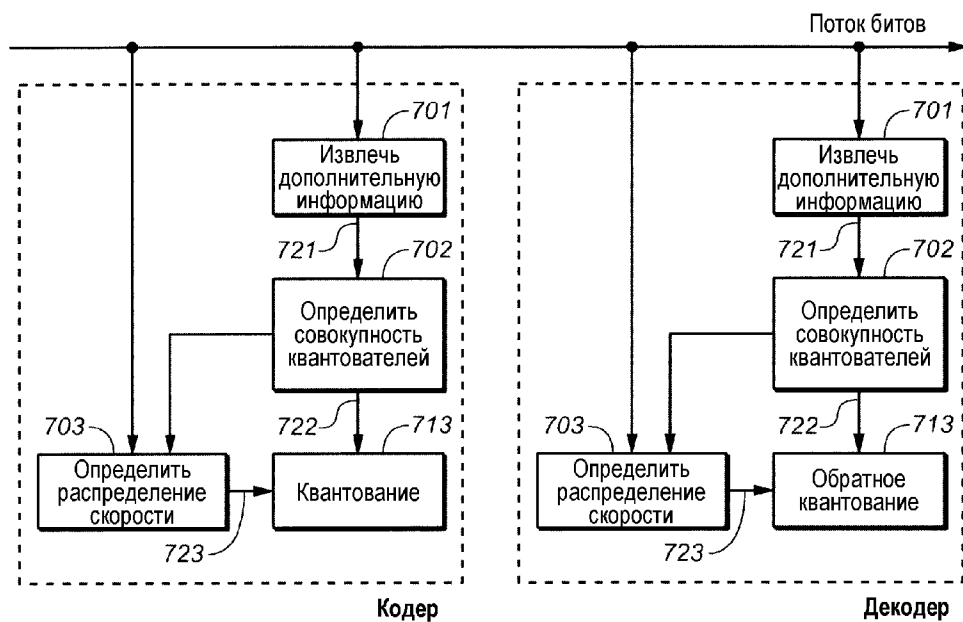
ФИГ. 24 б

23/25

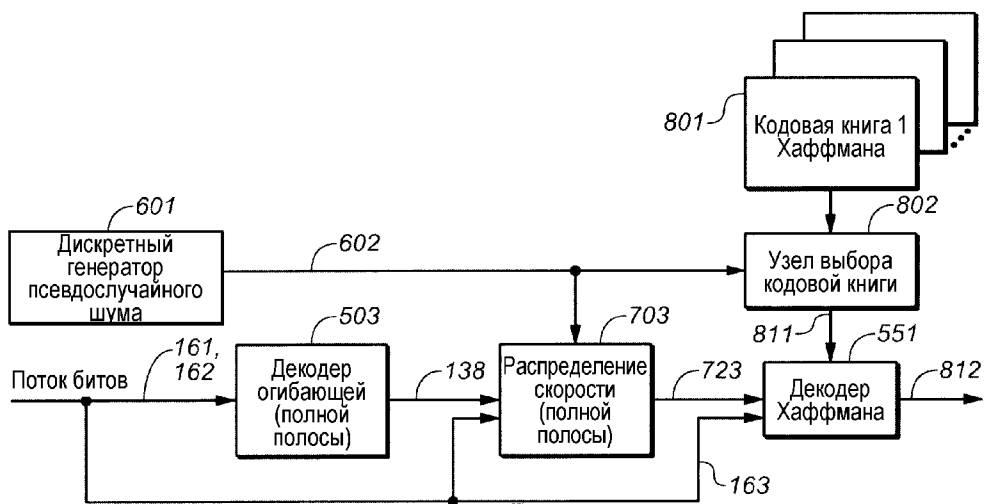


ФИГ. 24с

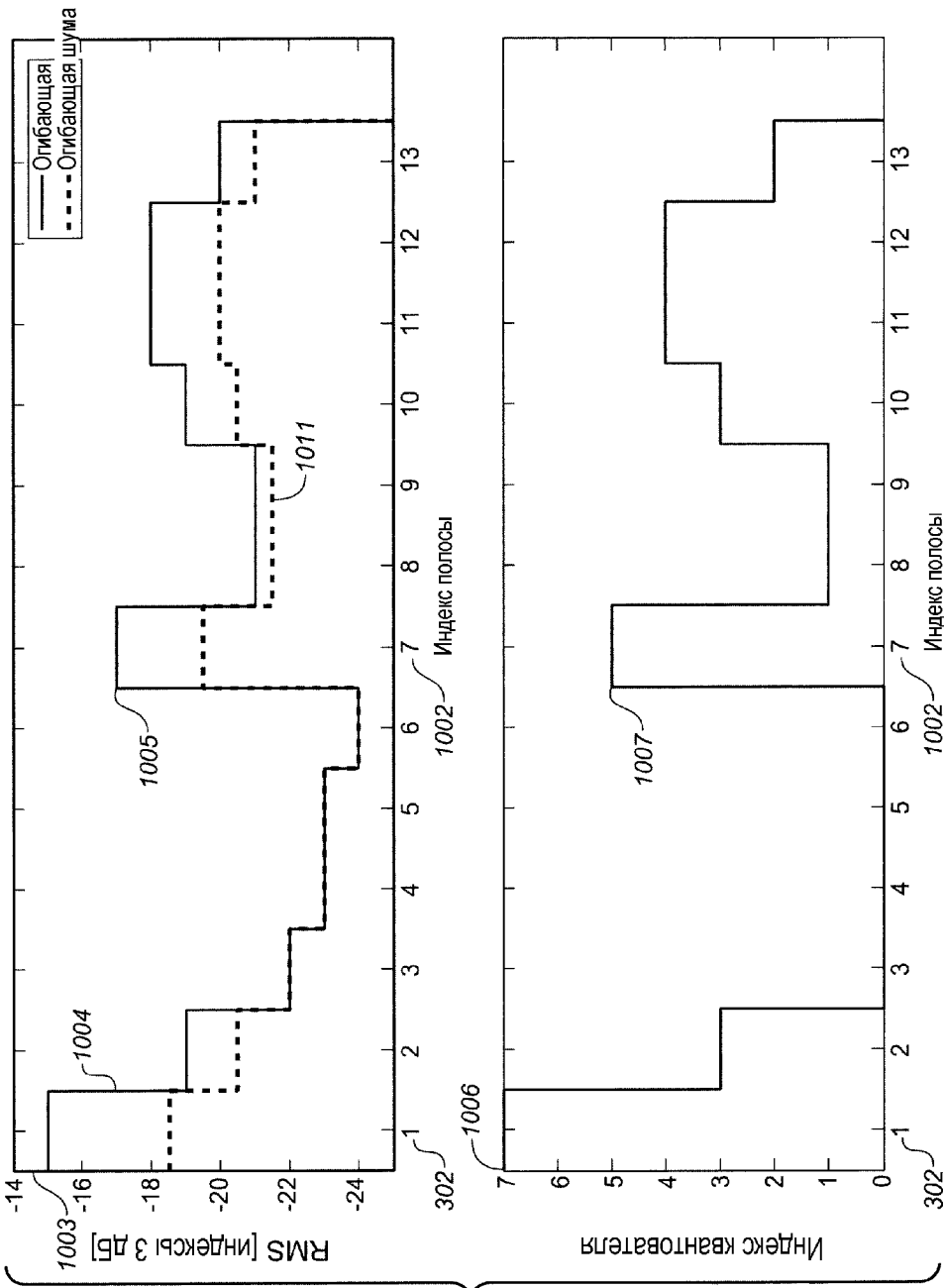
24/25



ФИГ. 25



ФИГ. 26



ФИГ. 27