



ФЕДЕРАЛЬНАЯ СЛУЖБА
ПО ИНТЕЛЛЕКТУАЛЬНОЙ СОБСТВЕННОСТИ

(12) ОПИСАНИЕ ИЗОБРЕТЕНИЯ К ПАТЕНТУ

(52) СПК
G10L 19/008 (2022.08)

(21)(22) Заявка: 2019122302, 16.07.2019

(24) Дата начала отсчета срока действия патента:
30.12.2008

Дата регистрации:
05.04.2023

Приоритет(ы):

(30) Конвенционный приоритет:
04.01.2008 SE 0800032-5;
24.05.2008 EP 08009530.0;
24.05.2008 US 61/055,978

Номер и дата приоритета первоначальной заявки,
из которой данная заявка выделена:
2015118725 04.01.2008

(43) Дата публикации заявки: 18.01.2021 Бюл. № 2

(45) Опубликовано: 05.04.2023 Бюл. № 10

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

(72) Автор(ы):

ХЕДЕЛИН, Пер, Хенрик (SE),
КАРЛССОН, Понтус, Ян (SE),
САМУЭЛЬССОН, Йонас, Лейф (SE),
ШУГ, Михель (DE)

(73) Патентообладатель(и):

ДОЛБИ ИНТЕРНЭШНЛ АБ (NL)

(56) Список документов, цитированных в отчете
о поиске: US 2004/0220805 A1, 04.11.2004. US
2006/0173675 A1, 03.08.2006. US 2005/0154597
A1, 14.07.2005. US 2007/0165869 A1, 19.07.2007.
RU 2158478 C2, 27.10.2000.

(54) АУДИОКОДЕР И ДЕКОДЕР

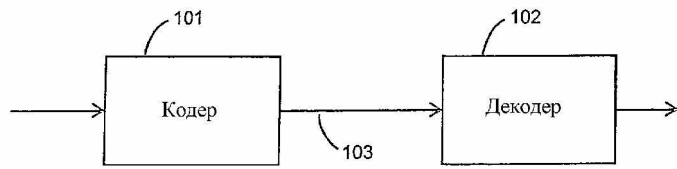
(57) Реферат:

Изобретение относится к области аудиообработки данных. Технический результат заключается в повышении точности кодирования как общих звуковых сигналов, так и речевых сигналов при низких скоростях передачи данных. Технический результат достигается за счет того, что создают кривые усиления MDCT-области для кадров коэффициентов MDCT на основе величин откликов, определенных из последовательности полиномов LPC, причем создание кривых

усиления в MDCT-области содержит отображение полиномов LPC в соответствующие кадры коэффициентов MDCT; применяют кривые усиления в MDCT-области к кадрам коэффициентов MDCT для получения кадров коэффициентов MDCT, скорректированных на усиление; и квантуют кадры коэффициентов MDCT, скорректированных на усиление. 7 н. и 9 з.п. ф-лы, 34 ил.

RU 2 7 9 3 7 2 5 C 2

RU 2 7 9 3 7 2 5 C 2



Фиг. 1

RU 2793725 C2

RU 2793725 C2



FEDERAL SERVICE
FOR INTELLECTUAL PROPERTY

(12) **ABSTRACT OF INVENTION**

(52) CPC
G10L 19/008 (2022.08)

(21)(22) Application: **2019122302, 16.07.2019**

(24) Effective date for property rights:
30.12.2008

Registration date:
05.04.2023

Priority:

(30) Convention priority:
04.01.2008 SE 0800032-5;
24.05.2008 EP 08009530.0;
24.05.2008 US 61/055,978

Number and date of priority of the initial application,
from which the given application is allocated:
2015118725 04.01.2008

(43) Application published: **18.01.2021 Bull. № 2**

(45) Date of publication: **05.04.2023 Bull. № 10**

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

(72) Inventor(s):

HEDELIN, Per, Henrik (SE),
CARLSSON, Pontus, Jan (SE),
SAMUELSSON, Jonas, Leif (SE),
SCHUG, Michael (DE)

(73) Proprietor(s):

DOLBY INTERNATIONAL AB (NL)

(54) **AUDIO CODER AND DECODER**

(57) Abstract:

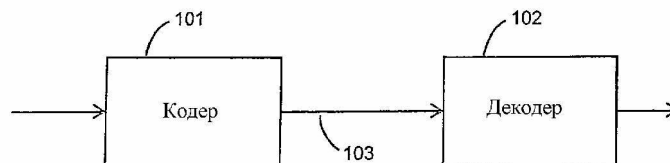
FIELD: audio data processing.

SUBSTANCE: technical result is achieved due to the fact that gain curves of the MDCT area for frames of MDCT coefficients are created based on the response values determined from the sequence of LPC polynomials, and the creation of gain curves in the MDCT area comprises mapping LPC polynomials to the corresponding frames of MDCT coefficients; gain

curves in the MDCT area are applied to the MDCT coefficient frames to obtain gain-corrected MDCT coefficient frames; and frames of gain-corrected MDCT coefficients are quantized.

EFFECT: increase in the coding accuracy of both general audio signals and speech signals at low data rates.

16 cl, 34 dwg



Фиг. 1

RU 2 793 725 C2

RU 2 793 725 C2

Область техники, к которой относится изобретение

Настоящее изобретение относится к кодированию аудиосигналов и, в частности, к кодированию любого аудиосигнала, не ограничиваясь речью, музыкой или их комбинацией.

5 Уровень техники изобретения

В предшествующем уровне техники существуют речевые кодеры, специально предназначенные для кодирования речевых сигналов, основанные на модели источника сигнала, то есть, системе человеческого голоса. Эти кодеры не могут обрабатывать произвольные аудиосигналы, такие как музыка или любой другой неречевой сигнал.
10 Кроме того, в предшествующем уровне техники существуют музыкальные кодеры, обычно называемые аудиокодерами, основывающие свое кодирование на предположениях о человеческой слуховой системе, а не на модели источника сигнала. Эти кодеры очень хорошо могут обрабатывать произвольные сигналы, хотя при низких скоростях передачи речевых сигналов специализированный речевой кодер дает
15 превосходное аудиокачество. Следовательно, на сегодня не существует общей структуры кодирования произвольных аудиосигналов, которая одинаково хороша как в качестве речевого кодера для речи, так и в качестве музыкального кодера для музыки, когда работа ведется при низких скоростях передачи данных.

Таким образом, существует потребность в улучшенном аудиокодере и декодере с
20 улучшенным аудиокачеством и/или пониженными скоростями передачи данных.

Сущность изобретения

Настоящее изобретение относится к эффективному кодированию произвольных аудиосигналов при уровне качества, равном или лучше, чем уровень качества системы, специально приспособленной к конкретному сигналу.

25 Настоящее изобретение обращено к алгоритмам аудиокодеков, содержащих как кодирование с линейным предсказанием (LPC), так и к части трансформного кодера, работающей с сигналами на основе обработки LPC.

Настоящее изобретение дополнительно относится к стратегии квантования, зависящей от размера преобразованного кадра данных. Дополнительно предлагается устройство
30 квантования, основанное на модели, с ограничением энтропии, использующее арифметическое кодирование. Кроме того, обеспечивается вставка случайных сдвигов в однородное скалярное устройство квантования. Изобретение дополнительно предлагает устройство квантования, основанное на модели, например, устройство квантования с ограничением энтропии (ECQ), использующее арифметическое
35 кодирование.

Настоящее изобретение дополнительно относится к эффективному кодированию масштабных коэффициентов в части трансформного кодирования аудиокодера, используя присутствие данных LPC.

40 Настоящее изобретение дополнительно относится к эффективному осуществлению использования битового накопителя в аудиокодере с переменным размером кадра.

Настоящее изобретение дополнительно относится к кодеру для кодирования аудиосигналов и создания битового потока передачи данных и к декодеру для декодирования битового потока передачи данных и созданию реконструированного аудиосигнала, перцептуально неотличимого от входного аудиосигнала.

45 Первый аспект настоящего изобретения относится к квантованию в трансформном кодере, которые применяется, например, при модифицированном дискретном косинусном преобразовании (MDCT). Предложенное устройство квантования предпочтительно квантует линии MDCT. Этот аспект применяется независимо от того,

использует ли кодер дополнительно анализ кодирования с линейным предсказанием (LPC) или дополнительное долгосрочное предсказание.

Настоящее изобретение обеспечивает систему кодирования аудиосигнала, содержащую блок линейного предсказания для фильтрации входного сигнала, основанный на адаптивном фильтре; блок преобразования для преобразования кадра отфильтрованного входного сигнала в область преобразования; и блок квантования для квантования сигнала в области преобразования. Блок квантования, основываясь на характеристиках входного сигнала, принимает решение о кодировании сигнала в области преобразования с помощью устройства квантования, основанного на модели, или устройства квантования, не основанного на модели. Предпочтительно, решение основывается на размере кадра, применяемом блоком преобразования. Однако, предусмотрены также и другие зависимые от входного сигнала критерии для переключения стратегии квантования, находящиеся в пределах объема настоящей заявки.

Другой важный аспект изобретения состоит в том, что устройство квантования может быть адаптивным. В частности, модель в устройстве квантования, основанном на модели, может быть адаптивной, чтобы регулировать входной аудиосигнал. Модель может, например, меняться во времени, например, в зависимости от характеристик входного сигнала. Это позволяет снизить искажение при квантовании и, таким образом, улучшить качество кодирования.

В соответствии с вариантом осуществления, предложенная стратегия квантования основывается на размере кадра. Предложено, чтобы модуль квантования, основываясь на размере кадра, применяемом блоком преобразования, мог принимать решение, кодировать ли сигнал в области преобразования с помощью устройства квантования, основанного на модели, или устройства квантования, не основанного на модели. Предпочтительно, блок квантования может быть выполнен с возможностью кодирования сигнала в области преобразования для кадра с размером кадра, меньшим, чем пороговое значение, посредством устройства квантования, основанного на модели, с ограничением энтропии. Квантование, основанное на модели, может выполняться на основе различных параметров. Большие кадры могут квантоваться, например, скалярным устройством квантования, например, с помощью энтропийного кодирования по способу Хаффмана, используемого, например, в кодеке AAC.

Система аудиокодирования может дополнительно содержать блок долгосрочного предсказания (LTP) для оценки кадра отфильтрованного входного сигнала, основываясь на реконструкции предыдущего сегмента отфильтрованного входного сигнала и сигнала в области преобразования блока объединения для объединения в области преобразования результата определения долгосрочного предсказания и преобразованного входного сигнала, чтобы создать сигнал в области преобразования, являющийся входным сигналом для блока квантования.

Переключение между различными способами квантования линий MDCT является другим аспектом предпочтительного варианта осуществления изобретения. Используя различные стратегии квантования для различных размеров преобразования, кодек может выполнять все квантование и кодирование в MDCT-области без необходимости иметь специальный речевой кодер временной области, работающий параллельно или последовательно с кодеком, работающим в области преобразования. Настоящее изобретение указывает, что для сигналов, подобных речи, где существует усиление LTP, сигнал предпочтительно кодируется, используя быстрого преобразование и устройство квантования, основанное на модели. Устройство квантования, основанное на модели,

в частности, пригодно для быстрого преобразования и обладает, как будет описано позже, преимуществами специального векторного устройства квантования (VQ) речи во временной области, все еще продолжая работать в MDCT-области, и без каких-либо требований, чтобы входной сигнал был речевым сигналом. Другими словами, когда
5 устройство квантования, основанное на модели, используется для сегментов быстрого преобразования в комбинации с LTP, эффективность специализированного кодера VQ речи, работающего во временной области, сохраняется без потери общности и без ухода из MDCT-области.

Кроме того, для более стационарных музыкальных сигналов предпочтительно
10 использовать преобразование относительно большого размера, которое обычно используется в аудиокодеках, и схему квантования, которая может пользоваться преимуществами, даваемыми разреженными спектральными линиями, различаемыми большим преобразованием. Поэтому настоящее изобретение указывает на использование этого вида схемы квантования для длинных преобразований.

Таким образом, переключение стратегии квантования как функция размера кадра
15 позволяет кодеку сохранять как свойства специализированного речевого кодека, так и свойства специализированного аудиокодека, просто посредством выбора размера преобразования. Это позволяет полностью избежать проблем, присущих системам предшествующего уровня техники, которые стремятся обрабатывать речевые и
20 аудиосигналы одинаково хорошо при низких скоростях, так как эти системы неизбежно сталкиваются с проблемами и трудностями эффективного объединения кодирования во временной области (речевой кодер) с кодированием в частотной области (аудиокодер).

В соответствии с другим аспектом изобретения, квантование использует адаптивные
25 размеры шага. Предпочтительно, размер(-ы) шага квантования для компонентов сигнала в области преобразования адаптируется/адаптируются, основываясь на параметрах линейного предсказания и/или долговременного предсказания. Размер(-ы) шага квантования может дополнительно быть выполнен с возможностью быть частотнозависимым. В вариантах осуществления изобретения размер шага квантования определяется, основываясь, по меньшей мере, на одном из следующего: полином
30 адаптивного фильтра, параметр управления скоростью кодирования, значение усиления при долгосрочном предсказании и дисперсия входного сигнала.

Предпочтительно, блок квантования содержит однородные скалярные устройства
квантования для квантования компонент сигнала в области преобразования. Каждое скалярное устройство квантования применяет к линии MDCT однородное квантование,
35 например, основанное на вероятностной модели. Вероятностная модель может лапласовой или гауссовой моделью или любой другой вероятностной моделью, пригодной для характеристик сигнала. Блок квантования может дополнительно вставлять в однородные скалярные устройства квантования случайный сдвиг. Вставка случайного сдвига обеспечивает однородным скалярным устройствам квантования
40 преимущества векторного квантования. В соответствии с вариантом осуществления, случайные сдвиги определяются, основываясь на оптимизации искажения при квантовании, предпочтительно в перцепционной области и/или при рассмотрении стоимости с точки зрения числа битов, требующихся для кодирования индексов квантования.

Блок квантования может дополнительно содержать арифметический кодер для
45 кодирования индексов квантования, созданных однородными скалярными устройствами квантования. Это позволяет достигнуть низкой скорости передачи данных, приближающейся к возможному минимуму, который задается энтропией сигнала.

Модуль квантования может дополнительно содержать устройство квантования остатка для квантования остатка сигнала квантования, являющегося результатом работы однородных скалярных устройств квантования, чтобы дополнительно уменьшить общее искажение. Устройство квантования остатка предпочтительно является векторным устройством с фиксированной частотой.

Многочисленные точки реконструкции квантования могут использоваться в блоке деквантования кодера и/или инверсном устройстве квантования в декодере. Например, точка реконструкции с минимальной среднеквадратичной ошибкой (MMSE) и/или центральная точка (средняя точка) реконструкции могут использоваться для реконструкции квантованного значения, основываясь на его индексе квантования. Точка реконструкции квантования может дополнительно основываться на динамической интерполяции между центральной точкой и точкой MMSE, возможно управляемыми характеристиками данных. Это позволяет управлять шумовой вставкой и избегать спектральных провалов из-за назначения линиям MDCT нулевого элемента квантования для низких скоростей передачи данных.

Перцепционное взвешивание в области преобразования предпочтительно применяется при определении искажения за счет квантования, чтобы придать различные веса конкретным частотным компонентам. Перцепционные веса могут эффективно быть получены из параметров линейного предсказания.

Другой независимый аспект изобретения имеет отношение к общей концепции использования сосуществования данных LPC and SCF (ScaleFactor). В трансформном кодере, например применяющем модифицированное дискретное косинусное преобразование (MDCT), масштабные факторы могут использоваться при квантовании, чтобы управлять размером шага квантования. На предшествующем уровне техники эти масштабные коэффициенты определяются из исходного сигнала, чтобы определить кривую маскирования. Теперь предлагается определить второй набор масштабных коэффициентов с помощью перцепционного фильтра или психоакустической модели, которая вычисляется из данных LPC. Это позволяет снизить стоимость передачи/хранения масштабных коэффициентов посредством передачи/хранения только отличия фактически применяемых масштабных коэффициентов от определяемых по LPC масштабных коэффициентов вместо передачи/сохранения реальных масштабных коэффициентов. Таким образом, в системе аудиокодирования, содержащей элементы кодирования речи, такие как, например, LPC, и элементы трансформного кодирования, такие как MDCT, настоящее изобретение уменьшает стоимость передачи информации масштабных коэффициентов, необходимой для трансформной части кодирования кодека, используя данные, предоставленные LPC. Следует отметить, что этот аспект независим от других аспектов предложенной системы аудиокодирования и может быть осуществлен также в другой системе аудиокодирования.

Например, кривая перцепционного маскирования может быть определена, основываясь на параметрах адаптивного фильтра. Линейное предсказание, основанное на втором наборе масштабных коэффициентов, может быть определено, основываясь на определенной кривой перцепционного маскирования. Сохраненная/переданная информация масштабных коэффициентов затем определяется на основе различия между масштабными коэффициентами, фактически используемыми при квантовании, и масштабными коэффициентами, вычисленными по кривой перцепционного маскирования, основанной на LPC. Это удаляет динамику и избыточность из сохраняемой/передаваемой информации, так чтобы было необходимо меньше битов для сохранения/передачи масштабных коэффициентов.

В случае, если когда LPC и MDCT работают не на одной и той же частоте кадров, то есть, имеют различные размеры кадров, масштабные коэффициенты, основанные на линейном предсказании, для кадров сигнала в области преобразования могут быть определены, основываясь на интерполированных параметрах линейного предсказания, чтобы соответствовать временному окну, охватываемому кадром MDCT.

Настоящее изобретение, следовательно, обеспечивает систему аудиокодирования, основанную на трансформном кодере, и содержит основополагающие модули предсказания и формирования из речевого кодера. Обладающая признаками изобретения система содержит блок линейного предсказания для фильтрации входного сигнала, основанный на адаптивном фильтре; блок преобразования для преобразования кадра отфильтрованного входного сигнала в область преобразования; блок квантования для квантования сигнала в области преобразования; блок определения масштабных коэффициентов для создания масштабных коэффициентов, основываясь на пороговой кривой маскирования, для использования в блоке квантования при квантовании сигнала в области преобразования; блок определения масштабных коэффициентов линейного предсказания для определения линейного предсказания, основываясь на масштабных коэффициентах, основанных на параметрах адаптивного фильтра; и кодер масштабных коэффициентов для кодирования разности масштабных коэффициентов, основанных на пороговой кривой маскирования, и масштабных коэффициентов, основанных на линейном предсказании. Кодируя разность между примененными масштабными коэффициентами и масштабными коэффициентами, которые могут быть определены в декодере, основанном на доступной информации линейного предсказания, эффективность кодирования и хранения может быть улучшена и сохранять/передавать потребуется совсем немного битов.

Другой независимый аспект изобретения, касающийся кодера, относится к обработке битового накопителя для кадров переменной длины. В системе аудиокодирования, которая может кодировать кадры переменной длины, битовый накопитель управляется посредством распределения доступных битов между кадрами. Учитывая разумную степень сложности индивидуальных кадров и битовый накопитель определенного размера, определенное отклонение от требуемой постоянной скорости передачи позволяет иметь лучшее общее качество без нарушения требований буфера, которые налагаются размером битового накопителя. Настоящее изобретение распространяет концепцию использования битового накопителя на управление битовым накопителем для общего аудиокодека с переменными размерами кадров. Система аудиокодирования может поэтому содержать блок управления битовым накопителем для определения числа битов, предоставляемых для кодирования кадра отфильтрованного сигнала, основываясь на длительности кадра и мере сложности кадра. Предпочтительно, блок управления битовым накопителем имеет разные уравнения управления для различных мер сложности кадров и/или различных размеров кадров. Меры сложности для различных размеров кадров могут быть нормализованы, так чтобы их можно было легче сравнивать. Чтобы управлять распределением бит для кодера с переменной частотой, блок управления битовым накопителем предпочтительно устанавливает в предоставленном алгоритме управления битами пониженный допустимый предел относительно среднего числа битов для наибольшего допустимого размера кадра.

Дополнительный аспект изобретения относится к обработке битового накопителя в кодере, используя основанное на модели устройство квантования, например, устройство квантования с ограничением энтропии (ECQ). Предлагается минимизировать изменение размера шага ECQ. Предлагается специальное уравнение управления,

связывающее размер шага устройства квантования со скоростью работы ЕСQ.

Адаптивный фильтр для фильтрации входного сигнала предпочтительно основывается на анализе кодирования с линейным предсказанием (LPC), содержащем фильтр LPC, создающий отбеленный входной сигнал. Параметры LPC для текущего кадра входных данных могут быть определены с помощью известных в технике алгоритмов. Блок определения параметров LPC может вычислить для кадра входных данных любое подходящее представление параметров LPC, такое как полиномы, передаточные функции, коэффициенты отражения, дискретные спектральные частоты и т.д. Конкретный тип представления параметров LPC, который используется для кодирования или другой обработки, зависит от соответствующих требований. Как известно специалистам в данной области техники, некоторые представления более пригодны для определенных операций, чем другие, и поэтому предпочтительны для выполнения этих операций. Блок линейного предсказания может воздействовать на первой длительности кадра, которая устанавливается равной, например, 20 мс. Фильтрация с линейным предсказанием может дополнительно работать на нелинейной частотной оси, чтобы выборочно подчеркивать определенные частотные диапазоны, такие как низкие частоты, по сравнению с другими частотами.

Преобразованием, применяемым к кадру отфильтрованного входного сигнала, предпочтительно является модифицированное дискретное косинусное преобразование (MDCT), работающее с переменной длительностью второго кадра. Система аудиокодирования может содержать блок управления последовательностью окон, определяющий для блока входного сигнала длительности кадров для перекрытия окон MDCT, минимизируя функцию стоимости кодирования, предпочтительно, упрощенную перцепционную энтропию, для всего блока входного сигнала, содержащего несколько кадров. Таким образом, получается оптимальная сегментация блока входного сигнала на окна MDCT, имеющие соответствующие длительности вторых кадров. Как следствие, предлагается структура кодирования в области преобразования, содержащая элементы речевого кодера, с кадром MDCT адаптивной длительности как единственным базовым блоком для всей обработки, кроме LPC. Поскольку длительности кадров MDCT могут принимать много различных значений, может быть найдена оптимальная последовательность и можно избежать резких изменений размера кадров, как это обычно происходит на предшествующем уровне техники, где применяются только малый размер окна и большой размер окна. Кроме того, нет необходимости в переходных преобразовательных окнах с резкими краями, используемых при некоторых подходах предшествующего уровня техники для перехода между малым и большим размерами окон.

Предпочтительно, длительности последовательных окон MDCT изменяются, самое большее, с коэффициентом два (2) и/или длительности окон MDCT являются диадическими значениями. Более конкретно, длительности окон MDCT могут быть диадическими частями блока входного сигнала. Последовательность окон MDCT поэтому ограничивается заданными последовательностями, которые просты для кодирования с помощью небольшого числа битов. Кроме того, последовательность окон имеет плавные переходы размеров кадров, исключая, тем самым, резкие изменения размеров кадров.

Блок управления последовательностью окон может быть дополнительно выполнен с возможностью результатов определения долгосрочного предсказания, созданных блоком долгосрочного предсказания, для кандидатов на продолжительность окон при поиске последовательности длительностей окон MDCT, которая минимизирует функцию

стоимости кодирования блока входного сигнала. В этом варианте осуществления цикл долгосрочного предсказания замыкается при определении длительностей окон MDCT, что приводит к улучшенной последовательности окон MDCT, применяемых для кодирования.

5 Система аудиокодирования может дополнительно содержать кодер LPC для рекурсивного кодирования с переменной частотой дискретных спектральных частот или других соответствующих представлений параметров LPC, создаваемых блоком линейного предсказания для хранения и/или передачи на декодер. В соответствии с вариантом осуществления, обеспечивается блок интерполяции линейного предсказания,
10 чтобы интерполировать параметры линейного предсказания, созданные с частотой, соответствующей длительности первого кадра, так чтобы соответствовать переменным длительностям кадров сигнала в области преобразования.

В соответствии с аспектом изобретения, система аудиокодирования может содержать блок перцепционного моделирования, изменяющий характеристику адаптивного
15 фильтра посредством линейной частотной модуляции и/или наклона полинома LPC, созданного блоком линейного предсказания для кадра LPC. Перцепционная модель, полученная посредством модификации характеристик адаптивного фильтра, может использоваться для многих целей в системе. Например, она может применяться в качестве функции перцепционного взвешивания при квантовании или долгосрочном
20 предсказании.

Другой аспект изобретения относится к долгосрочному предсказанию (LTP), в частности, к долгосрочному предсказанию в MDCT-области, адаптированному LTP кадра MDCT и поиску LTP со взвешенным MDCT. Эти аспекты применимы независимо от того, присутствует ли анализ LPC в восходящем потоке данных трансформного
25 кодера.

В соответствии с вариантом осуществления, система аудиокодирования дополнительно содержит блок инверсного квантования и инверсного преобразования для создания реконструкции во временной области кадра отфильтрованного входного сигнала. Дополнительно может обеспечиваться буфер долгосрочного предсказания
30 для хранения реконструкций во временной области предыдущих кадров отфильтрованного входного сигнала. Эти блоки могут быть организованы в цикл с обратной связью от блока квантования к блоку извлечения долгосрочного предсказания, который ищет в буфере долгосрочного предсказания реконструированный сегмент, наилучшим образом совпадающий с текущим кадром отфильтрованного входного
35 сигнала. Кроме того, может быть обеспечен блок определения усиления долгосрочного предсказания, который регулирует усиление сегмента, выбранного из буфера долгосрочного предсказания, так чтобы он наилучшим образом совпадал с текущим кадром. Предпочтительно, результат определения долгосрочного предсказания вычитается из преобразованного входного сигнала в области преобразования. Поэтому
40 может обеспечиваться второй блок преобразования для преобразования выбранного сегмента в область преобразования. Цикл долгосрочного предсказания может дополнительно содержать добавление результата определения долгосрочного предсказания в области преобразования к сигналу обратной связи после инверсного квантования и перед инверсным преобразованием во временную область. Таким
45 образом, может использоваться схема обратного адаптивного долгосрочного предсказания, которая предсказывает в области преобразования текущий кадр отфильтрованного входного сигнала, основываясь на предыдущих кадрах. Для большей эффективности, схема долгосрочного предсказания может дополнительно быть

адаптирована различными способами, как изложено ниже для некоторых примеров.

В соответствии с вариантом осуществления, блок долгосрочного предсказания содержит устройство выделения долгосрочного предсказания для определения значения задержки, указывающее реконструированный сегмент отфильтрованного сигнала, наилучшим образом соответствующий текущему кадру отфильтрованного сигнала. Устройство определения усиления при долгосрочном предсказании может определить значение усиления, применяемого к сигналу выбранного сегмента отфильтрованного сигнала. Предпочтительно, значение задержки и значение усиления определяются так, чтобы минимизировать критерий искажения, относящийся к разности в перцепционной области между оценкой долгосрочного предсказания и преобразованным входным сигналом. Модифицированный полином линейного предсказания может применяться в качестве кривой усиления выравнивания в MDCT-области при минимизации критерия искажения.

Блок долгосрочного предсказания может содержать блок преобразования для преобразования реконструированных сегментов из буфера LTP в область преобразования. Для эффективного осуществления MDCT-преобразования такое преобразование предпочтительно должно быть дискретным косинусным преобразованием типа-IV.

Другой аспект изобретения относится к аудиодекодеру для декодирования битового потока, созданного с помощью вариантов осуществления упомянутого выше кодера. Декодер, соответствующий варианту осуществления, содержит блок деквантования для деквантования кадра входного битового потока, основываясь на масштабных коэффициентах; блок инверсного преобразования для инверсного преобразования сигнал в области преобразования; блок линейного предсказания для фильтрации инверсно преобразованного сигнала в области преобразования; и блок декодирования масштабных коэффициентов для создания масштабных коэффициентов, используемых при деквантовании, основываясь на полученной дельта-информации масштабных коэффициентов, которая кодирует разность между масштабными коэффициентами, применяемыми в кодере, и масштабными коэффициентами, созданными, основываясь на параметрах адаптивного фильтра. Декодер может дополнительно содержать блок определения масштабных коэффициентов для создания масштабных коэффициентов, основываясь на пороговой кривой маскирования, получаемой из параметров линейного предсказания для текущего кадра. Блок декодирования масштабных коэффициентов может объединить полученную дельта-информацию масштабных коэффициентов с созданным линейным предсказанием, основываясь на масштабных коэффициентах, чтобы создать масштабные коэффициенты для ввода в блок деквантования.

Декодер, соответствующий другому варианту осуществления, содержит блок деквантования, основанный на модели, для деквантования кадра входного битового потока; блок инверсного преобразования для инверсного преобразования сигнала в области преобразования; и блок линейного предсказания для фильтрации инверсно преобразованного сигнала в области преобразования. Блок деквантования может содержать устройство деквантования, основанное на модели, и устройство деквантования, не основанное на модели.

Предпочтительно, блок деквантования содержит, по меньшей мере, одну адаптивную вероятностную модель. Блок деквантования может быть выполнен с возможностью адаптации деквантования как функции характеристик переданного сигнала.

Блок деквантования может дополнительно принимать решение по стратегии деквантования, основываясь на данных управления для декодированного кадра.

Предпочтительно, данные управления деквантованием принимаются вместе с битовым потоком или получаются из принятых данных. Например, блок деквантования принимает решение по стратегии деквантования, основываясь на размере преобразования кадра.

5 В соответствии с другим аспектом, блок деквантования содержит точки адаптивной реконструкции. Блок деквантования может содержать однородные скалярные устройства деквантования, выполненные с возможностью использования двух точек реконструкции деквантования на интервале квантования, в частности, средней точки и точки реконструкции с MMSE.

10 В соответствии с вариантом осуществления, блок деквантования использует устройство квантования, основанное на модели, в комбинации с арифметическим кодированием.

Кроме того, декодер может содержать многие из аспектов, раскрытых выше для кодера. В целом, декодер будет отражать операции кодера, хотя некоторые операции
15 выполняются только в кодере и не будут иметь никаких соответствующих компонентов в декодере. Таким образом, то, что описано для кодера, должно считаться применимым также и для декодера, если не указывается иначе.

Упомянутые выше аспекты изобретения могут быть осуществлены как устройство, совокупность устройств, способ или компьютерная программа, работающая на
20 программируемом устройстве. Аспекты изобретения могут дополнительно быть осуществлены в сигналах, структурах данных и битовых потоках.

Таким образом, заявка дополнительно раскрывает способ аудиокодирования и способ аудиодекодирования. Пример способа аудиокодирования содержит этапы, на
25 которых: фильтруют входной сигнал на основе адаптивного фильтра; преобразуют кадр отфильтрованного входного сигнала в область преобразования; квантуют сигнал в области преобразования; создают масштабные коэффициенты, основываясь на пороговой кривой маскирования, для использования в блоке квантования при квантовании сигнала в области преобразования; определяют основанные на линейном предсказании масштабные коэффициенты, используя для этого параметры адаптивного
30 фильтра; и кодируют разность между масштабными коэффициентами, основанными на пороговой кривой маскирования, и масштабными коэффициентами, основанными на линейном предсказании.

Другой способ аудиокодирования содержит этапы, на которых: фильтруют входной сигнал, основываясь на адаптивном фильтре; преобразуют кадр отфильтрованного
35 входного сигнала в область преобразования; и квантуют сигнал в области преобразования; в котором блок квантования, основываясь на характеристиках входного сигнала, принимает решение кодировать сигнал в области преобразования с помощью устройства квантования, основанного на модели, или устройства квантования, не основанного на модели.

40 Пример способа аудиодекодирования содержит этапы, на которых: деквантуют кадр входного битового потока, основываясь на масштабных коэффициентах; инверсно преобразуют сигнал в области преобразования; фильтруют с линейным предсказанием инверсно преобразованный сигнал в области преобразования; определяют вторые масштабные коэффициенты, основываясь на параметрах адаптивного фильтра; и
45 создают масштабные коэффициенты, используемые при деквантовании, основываясь на полученной информации о разности масштабных коэффициентов и определенных вторых масштабных коэффициентах.

Другой способ аудиокодирования содержит этапы, на которых: деквантуют кадр

входного битового потока; инверсно преобразуют сигнал в области преобразования; и фильтруют с линейным предсказанием инверсно преобразованный сигнал в области преобразования; в котором деквантование использует устройство квантования, основанное на модели, и устройство квантования, не основанное на модели.

5 Приведенное выше является лишь примерами предпочтительных способов аудиокодирования/декодирования и компьютерных программ, которые предлагаются настоящей заявкой и которые специалист в данной области техники может получить из последующего описания примеров вариантов осуществления.

Краткое описание чертежей

10 Настоящее изобретение будет теперь описано посредством иллюстративных примеров, не ограничивающих объем или сущность изобретения, со ссылкой на сопроводительные чертежи, на которых:

Фиг. 1 - предпочтительный вариант осуществления кодера и декодера в соответствии с настоящим изобретением;

15 Фиг. 2 - более подробное представление кодера и декодера в соответствии с настоящим изобретением;

Фиг. 3 - другой вариант осуществления кодера, соответствующего настоящему изобретению;

20 Фиг. 4 - предпочтительный вариант осуществления кодера, соответствующего настоящему изобретению;

Фиг. 5 - предпочтительный вариант осуществления декодера, соответствующего настоящему изобретению;

Фиг. 6 - предпочтительный вариант осуществления кодирования и декодирования линий MDCT в соответствии с настоящим изобретением;

25 Фиг. 7 - предпочтительный вариант осуществления кодера и декодера и примеры соответствующих данных управления, передаваемых от одного к другому в соответствии с настоящим изобретением;

Фиг. 7a - другой пример аспектов кодера в соответствии с вариантом осуществления изобретения;

30 Фиг. 8 - пример последовательности окон и отношения между данными LPC и данными MDCT в соответствии с вариантом осуществления настоящего изобретения;

Фиг. 9 - объединение данных масштабных коэффициентов и данных LPC в соответствии с настоящим изобретением;

35 Фиг. 9a - другой вариант осуществления объединения данных масштабных коэффициентов и данных LPC в соответствии с настоящим изобретением;

Фиг. 9b - другая упрощенная блок-схема кодера и декодера в соответствии с настоящим изобретением;

Фиг. 10 - предпочтительный вариант осуществления перевода полиномов LPC в кривую усиления MDCT в соответствии с настоящим изобретением;

40 Фиг. 11 - предпочтительный вариант осуществления отображения параметров LPC с постоянной частотой обновления в данные последовательности окон с адаптивным MDCT в соответствии с настоящим изобретением;

Фиг. 12 - предпочтительный вариант осуществления вычисления адаптации перцепционного взвешивающего фильтра, основываясь на размере преобразования и типе устройства квантования в соответствии с настоящим изобретением;

45 Фиг. 13 - предпочтительный вариант осуществления адаптации устройства квантования в зависимости от размера кадра в соответствии с настоящим изобретением;

Фиг. 14 - предпочтительный вариант осуществления адаптации устройства

квантования в зависимости от размера кадра в соответствии с настоящим изобретением;

Фиг. 15 - предпочтительный вариант осуществления адаптации размера шага квантования как функции LPC и данных LTP в соответствии с настоящим изобретением;

5 Фиг. 15a – вывод дельта-кривой из параметров LPC и LTP с помощью блока дельта-адаптации;

Фиг. 16 - предпочтительный вариант осуществления устройства квантования, основанного на модели, использующего случайные сдвиги, соответствующего настоящему изобретению;

10 Фиг. 17 - предпочтительный вариант осуществления устройства квантования, основанного на модели, в соответствии с настоящим изобретением;

Фиг. 17a - другой предпочтительный вариант осуществления устройства квантования, основанного на модели, в соответствии с настоящим изобретением;

Фиг. 17b – схематичное представление декодера 2150, основанного на модели, для линий MDCT в соответствии с вариантом осуществления изобретения;

15 Фиг. 17c - схематичное представление аспектов предварительной обработки устройства квантования в соответствии с вариантом осуществления изобретения;

Фиг. 17d – схематичное представление аспектов вычисления размера шага в соответствии с вариантом осуществления изобретения;

20 Фиг. 17e – схематичное представление кодера, основанного на модели, с ограничением энтропии в соответствии с вариантом осуществления изобретения;

Фиг. 17f – схематичное представление работы однородного скалярного устройства квантования (USQ) в соответствии с вариантом осуществления изобретения;

Фиг. 17g – схематичное представление вычислений вероятности в соответствии с вариантом осуществления изобретения;

25 Фиг. 17h – схематичное представление процесс деквантования в соответствии с вариантом осуществления изобретения;

Фиг. 18 - предпочтительный вариант осуществления управления битовым накопителем в соответствии с настоящим изобретением;

Фиг. 18a – базовая концепция управления битовым накопителем;

30 Фиг. 18b – концепция управления битовым накопителем для переменных размеров кадра в соответствии с настоящим изобретением;

Фиг. 18c - пример кривой управления для управления битовым накопителем в соответствии с вариантом осуществления;

35 Фиг. 19 - предпочтительный вариант осуществления инверсного устройства квантования, использующего различные точки реконструкции в соответствии с настоящим изобретением.

Описание предпочтительных вариантов осуществления

Описанные ниже варианты осуществления являются просто иллюстративными примерами принципов настоящего изобретения для аудиокодера и декодера.

40 Подразумевается, что модификации и изменения схем и подробностей, описанных здесь, будут очевидны другим специалистам в данной области техники. Намерение, поэтому, заключается в том, чтобы ограничиться только объемом пунктов сопроводительной формулы изобретения, а не конкретными подробностями, представленными посредством описания и объяснения представленных здесь вариантов осуществления. Подобные
45 компоненты вариантов осуществления обозначаются подобными ссылочными номерами.

На фиг. 1 визуальны представлены кодер 101 и декодер 102. Кодер 101 получает входной сигнал во временной области и создает битовый поток 103, впоследствии

посылаемый на декодер 102. Декодер 102 создает выходной колебательный сигнал, основываясь на принятом битовом потоке 103. Выходной сигнал психоакустически схож с первоначальным входным сигналом. На фиг. 2 показан предпочтительный вариант осуществления кодера 200 и декодера 210. Входной сигнал кодера 200 проходит
5 через модуль 201 LPC (кодирования с линейным предсказанием), создающий отбеленный остаточный сигнал для кадра LPC, имеющего первую длительность кадра, и соответствующие параметры линейного предсказания. Дополнительно, в модуль 201 LPC может быть включена нормализация усиления. Остаточный сигнал от LPC преобразуется в частотную область с помощью модуля 202 MDCT (модифицированного
10 дискретного косинусного преобразования), работающего на второй переменной длительности кадра. В кодере 200, показанном на фиг. 2, содержится модуль 205 LTP (долгосрочного предсказания). LTP будет с подробностями раскрыто в дополнительном варианте осуществления настоящего изобретения. Линии MDCT подвергаются процессу 203 квантования, а также процессу 204 деквантования, чтобы предоставить буферу LTP копию декодированного выходного сигнала, когда он будет доступен декодеру 210.
15 Из-за искажения при квантовании эту копию называют реконструкцией соответствующего входного сигнала. В нижней части фиг. 2 изображен декодер 210. Декодер 210 получает квантованные линии MDCT, производит процесс 211 их деквантования, добавляет вклад от модуля 214 LTP и производит процесс 212 инверсного преобразования MDCT с последующим синтезом фильтром 213 LPC.
20

Важным аспектом описанного выше варианта осуществления является то, что кадр MDCT является единственным базовым блоком для кодирования, хотя LPC имеет свой собственный (и в одном варианте осуществления постоянный) размер кадра и параметры LPC также кодируются. Вариант осуществления начинается с кодера преобразования
25 и вводит основополагающие модули предсказания и формирования из речевого кодера. Как будет обсуждаться позже, размер кадра MDCT является переменным и адаптируется к блоку входного сигнала, определяя оптимальную последовательность окон MDCT для всего блока посредством минимизации упрощенной функции стоимости перцепционной энтропии. Это позволяет осуществлять масштабирование, чтобы
30 поддержать оптимальное управление по времени/частоте. Дополнительно, предложенная унифицированная структура избегает переключаемых или наслаивающихся комбинаций различных парадигм кодирования.

На фиг. 3 части кодера 300 описаны схематично более подробно. Отбеленный сигнал в качестве выходного сигнала модуля 201 LPC кодера, показанного на фиг. 2, является
35 входным сигналом блока 302 фильтров MDCT. Анализ MDCT, как вариант, может быть анализом MDCT с нелинейной шкалой времени, что гарантирует, что шаг сигнала (если сигнал является периодическим со строго определенным шагом) является постоянным в окне преобразования MDCT.

На фиг. 3 модуль 310 LTP представлен более подробно. Он содержит буфер 311 LTP,
40 хранящий реконструированные выборки во временной области предыдущих сегментов выходного сигнала. Устройство 312 выделения LTP находит сегмент с наилучшим соответствием в буфере 311 LTP для заданного текущего входного сегмента. Блоком 313 усиления к этому сегменту прикладывается подходящее значение усиления перед тем, как он будет вычитаться из сегмента, в настоящее время являющегося входным
45 для устройства 303 квантования. Очевидно, что для выполнения вычитания перед квантованием, устройство 312 выделения LTP также преобразует выбранный сегмент сигнала в MDCT-область. Устройство 312 выделения LTP ищет наилучшие значения усиления и задержки, которые минимизируют функцию ошибки в перцепционной области

при объединении реконструированного предыдущего сегмента выходного сигнала с преобразованным входным кадром MDCT-области. Например, оптимизируется функция среднеквадратичной ошибки (MSE) между преобразованным реконструированным сегментом модуля 310 LTP и преобразованным входным кадром (то есть, остаточный сигнал после вычитания). Эта оптимизация может быть выполнена в перцепционной области, где частотные компоненты (то есть, линии MDCT) взвешиваются согласно их перцепционной важности. Модуль 310 LTP работает в блоках кадров MDCT и кодер 300 считает один остаток кадра MDCT за один раз, например, для квантования в модуле 303 квантования. Поиск задержки и усиления могут выполняться в перцепционной области. Как вариант, LTP может быть селективным по частоте, то есть, адаптировать усиление и/или задержку в зависимости от частоты. Показаны блок 304 инверсного квантования 304 и блок 306 инверсного MDCT. MDCT может иметь нелинейную шкалу времени, как объясняется позже.

На фиг. 4 показан другой вариант осуществления кодера 400. В дополнение к фиг. 3, для ясности введен анализ 401 LPC. Показано преобразование 414 DCT-IV, используемое для преобразования выбранного сегмента сигнала в MDCT-область. Дополнительно показаны несколько способов вычисления минимальной ошибки для выбора сегмента LTP. В дополнение к минимизации остаточного сигнала, как показано на фиг. 4, (идентифицируется как LTP2 на фиг. 4) показана минимизация разности между преобразованным входным сигналом и деквантованным сигналом MDCT-области перед инверсным преобразованием в реконструированный сигнал временной области для хранения в буфере 411 LTP (обозначен как LTP3). Минимизация этой функции MSE будет направлять вклад LTP к оптимальному (насколько возможно) подобию преобразованного входного сигнала и реконструированного входного сигнала для хранения в буфере 411 LTP. Другая альтернативная функция ошибки (обозначенная как LTP1) основана на разности этих сигналов во временной области. В этом случае, MSE между отфильтрованным входным кадром LPC и соответствующей реконструкцией во временной области в буфере 411 LTP минимизируется. MSE предпочтительно вычисляется, основываясь на размере кадра MDCT, который может отличаться от размера кадра LPC. Дополнительно, устройство квантования и блоки деквантования заменяются блоком 403 спектрального кодирования и блоками 404 спектрального декодирования ("Spec enc" и "Spec dec"), которые может содержать дополнительные модули, помимо модулей квантования, как в общем виде указано на фиг. 6. И снова, MDCT и инверсное MDCT могут иметь нелинейную шкалу времени (WMDCT, IWMDCT).

На фиг. 5 показан предложенный декодер 500. Данные спектра из принятого битового потока инверсно квантуются 511 и складываются с вкладом LTP, обеспечиваемым устройством выделения LTP из буфера 515 LTP. Также показаны устройство 516 выделения LTP и блок 517 усиления LTP в декодере 500. Суммированные линии MDCT синтезируются во временной области блоком синтеза MDCT и сигнал во временной области спектрально формируется фильтром 513 синтеза LPC.

На фиг. 6 блоки 403, 404, "Spec dec" и "Spec enc", показанные на фиг. 4, описаны более подробно. Блок 603 "Spec enc", показанный в правой части чертежа, содержит в варианте осуществления модуль 610 анализа гармонического предсказания, модуль 611 анализа TNS (временного ограничения шума), сопровождаемый модулем 612 масштабирования масштабных коэффициентов для линий MDCT, и, наконец, квантование и кодирование линий в модуле 613 линий Enc. Блок 604 "Spec Dec" декодера, показанный в левой части чертежа, выполняет инверсный процесс, то есть, принятые линии MDCT деквантуются в модуле 620 линий Dec и масштабирование уничтожается модулем 621 масштабирования

масштабных коэффициентов (SCF). Применяются синтез 622 TNS и синтез 623 гармонического предсказания.

На фиг. 7 показано очень общее представление системы кодирования, соответствующей изобретению. Кодер, как пример, получает входной сигнал и создает битовый поток, содержащий, среди других данных:

- квантованные линии MDCT;
- масштабные коэффициенты;
- полиномиальное представление LPC;
- энергию сегмента сигнала (например, дисперсия сигнала);
- последовательность окон;
- данные LTP.

Декодер, соответствующий варианту осуществления, считывает предоставленный битовый поток и создает выходной аудиосигнал, психоакустически схожий с исходным сигналом.

На фиг. 7а представлены другие аспекты кодера 700, соответствующего варианту осуществления изобретения. Кодер 700 содержит модуль 701 LPC, модуль 704 MDCT, модуль 705 LTP (показан только упрощенно), модуль 703 квантования и модуль 704 инверсного квантования для возвращения реконструированных сигналов обратно в модуль 705 LTP. Дополнительно обеспечиваются модуль 750 определения шага для определения шага входного сигнала и модуль 751 определения последовательности окон для определения оптимальной последовательности окон MDCT для блока входного сигнала большего размера (например, 1 секунда). В этом варианте осуществления последовательность окон MDCT определяется, основываясь на подходе с разомкнутым циклом, при котором определяется последовательность кандидатов на размер окна MDCT, что минимизирует функцию стоимости кодирования, например, упрощенную перцепционную энтропию. Вклад модуля 705 LTP в функцию стоимости кодирования, которая минимизируется модулем 751 определения последовательности окон, как вариант, можно учитывать при поиске оптимальной последовательности окон MDCT. Предпочтительно, для каждого определенного кандидата на размер окна определяется наилучший вклад долговременного предсказания в кадр MDCT, соответствующий кандидату на размер окна, и определяется соответствующая стоимость кодирования. В целом, короткие размеры кадра MDCT более подходят для ввода речи, тогда как окна длительного преобразования, имеющие прекрасную спектральную разрешающую способность, предпочтительны для аудиосигналов.

Перцепционные веса или функция перцепционного взвешивания определяются на основе параметров LPC, вычисляемых модулем 701 LPC, который ниже будет объяснен более подробно. Перцепционные веса подаются на модуль 705 LTP и на модуль 703 квантования 703, оба работающие в MDCT-области, для взвешивания ошибок или вклада искажений частотных компонент в соответствии с их соответствующей перцепционной важностью. На фиг. 7а дополнительно показано, какие параметры кодирования передаются на декодер, предпочтительно, соответствующей схемой кодирования, как будет обсуждаться позже.

Далее будет обсуждаться сосуществование данных LPC и MDCT и эмуляция эффекта LPC в MDCT, как для противодействия, так и для пропуска фактической фильтрации.

В соответствии с вариантом осуществления, модуль LP фильтрует входной сигнал так, что спектральная форма сигнала удаляется и последующий выходной сигнал модуля LP является спектрально плоским сигналом. Это предпочтительно, например, для работы LTP. Однако, другие части кодека, работающие со спектрально плоским

сигналом, могут получать выгоду от знания того, какая спектральная форма исходного сигнала предшествовала фильтрации LP. Поскольку модули кодера после фильтрации работают с преобразованием MDCT спектрально плоского сигнала, настоящее изобретение указывает, что спектральная форма первоначального сигнала до
5 фильтрации с LP может, если нужно, быть повторно наложена на MDCT-представление спектрально плоского сигнала, отображая передаточную функцию используемого LP-фильтра (то есть, огибающую спектра исходного сигнала) на кривой усиления или кривой выравнивания, которая применяется на элементах разрешения по частоте MDCT-представления спектрально плоского сигнала. Напротив, модуль LP может исключить
10 фактическую фильтрацию и определять только передаточную функцию, которая впоследствии отображается на кривой усиления, которая может быть наложена на MDCT-представление сигнала, тем самым исключая потребность в фильтрации во временной области входного сигнала.

Одним из явных аспектов вариантов осуществления настоящего изобретения является то, что трансформный кодер на MDCT-основе работает, используя гибкую сегментацию
15 окон, на отбеленном сигнале LPC. Это показано на фиг. 8, где приведен пример последовательности окон MDCT, наряду с работой с окнами LPC. Следовательно, как ясно из чертежа, LPC работает с постоянным размером кадров (например, 20 мс), тогда как MDCT работает с переменной последовательностью окон (например, 4-128 мс).
20 Это позволяет независимо выбирать оптимальную длительность окна для LPC и оптимальную последовательность окон для MDCT.

Фиг. 8 дополнительно показывает связь между данными LPC, в частности, параметрами LPC, созданными при первой частоте кадров, и данными MDCT, в частности, линиями MDCT, созданные при второй переменной частоте. Направленные
25 вниз стрелки на чертеже символизируют данные LPC, которые интерполируются между кадрами LPC (окружности), так чтобы совпадать с соответствующими кадрами MDCT. Например, созданная с использованием LPC функция перцепционного взвешивания интерполируется для временных случаев, как они определяются последовательностью окон MDCT. Стрелки, направленные вверх, символизируют данные уточнения (то есть,
30 данные управления), используемые для кодирования линий MDCT. Для кадров AAC эти данные обычно являются масштабными коэффициентами, а для кадров ECQ данные обычно являются данными коррекции дисперсии и т.д. Сплошные линии относительно пунктирных линий представляют, какие данные являются "самыми важными" данными для кодирования линий MDCT для определенного устройства квантования. Двойные
35 направленные вниз стрелки символизируют спектральные линии кодека.

Сосуществование LPC и данных MDCT в кодере может использоваться, например, для уменьшения потребности в битах при кодировании масштабных коэффициентов MDCT, принимая во внимание кривую перцепционного маскирования, определяемую из параметров LPC. Дополнительно, выведенное из LPC перцепционное взвешивание
40 может использоваться при определении искажения квантования. Как показано на чертеже и как будет обсуждено ниже, устройство квантования работает в двух режимах и создает два типа кадров (кадры ECQ и кадры AAC) в зависимости от размера кадров принятых данных, то есть, соответствующие размеру кадра или окна MDCT.

На фиг. 11 показан предпочтительный вариант осуществления отображения параметров LPC с постоянной частотой в данные адаптивной последовательности окон MDCT. Модуль 1100 отображения LPC принимает параметры LPC в соответствии с частотой обновления LPC. Кроме того, модуль 1100 отображения LPC принимает
45 информацию о последовательности окон MDCT. Затем он создает отображение LPC-

в-MDCT, например, отображение психоакустических данных, основанных на LPC, в соответствующие кадры MDCT, созданные с переменной частотой кадров MDCT. Например, модуль отображения LPC интерполирует полиномы LPC или сопутствующие данные для временных случаев, соответствующих кадрам MDCT, для использования, например, в качестве перцепционных весов в модуле LTP или устройстве квантования.

Теперь, специфика перцепционной модели, основанной на LPC, обсуждается со ссылкой на фиг. 9. Модуль 901 LPC находится в варианте осуществления настоящего изобретения, выполненном с возможностью создания белого выходного сигнала, используя линейное предсказание, например, порядка 16 для сигнала с частотой выборки 16 кГц. Например, выходной сигнал модуля 201 LPC на фиг. 2 является остаточным после определения и фильтрации параметров LPC. Определенный полином $A(z)$ LPC, как схематично показано внизу слева на фиг. 9, может подвергаться линейной частотной модуляции с коэффициентом расширения ширины полосы, а также наклоняться, в одной из реализаций изобретения, изменяя первый коэффициент отражения соответствующего полинома LPC. Линейная частотная модуляция расширяет ширину полосы пиков в передаточной функции LPC, перемещая полюса полинома внутрь единичной окружности, приводя, таким образом, в результате к более сглаженным пикам. Наклон позволяет делать передаточную функцию LPC более плоской, чтобы балансировать влияние нижних и верхних частот. Эти модификации стремятся создать кривую перцепционного маскирования $A'(z)$ из определенных параметров LPC, которые будут доступны как на стороне кодера, так и на стороне декодера системы. Подробности манипуляции полиномом LPC представлены ниже на фиг. 12.

Кодирование MDCT, применяемое к остатку LPC, имеет в одной из реализаций изобретения масштабные коэффициенты для управления разрешающей способностью устройства квантования или размерами шага квантования (и, таким образом, шумом, вносимым квантованием). Эти масштабные коэффициенты определяются модулем 960 определения масштабных коэффициентов для первоначального входного сигнала. Например, масштабные коэффициенты получаются из пороговой кривой перцепционного маскирования, определенной из первоначального сигнала. В варианте осуществления отдельное преобразование частот (имеющее, возможно, различную разрешающую способность по частоте) может использоваться для определения пороговой кривой маскирования, но это не всегда необходимо. Альтернативно, пороговая кривая маскирования определяется из линий MDCT, созданных модулем преобразования. В нижней правой части на фиг. 9 схематично показаны масштабные коэффициенты, созданные модулем 960 определения масштабных коэффициентов для управления квантованием, так чтобы внесенный шум квантования ограничивался неслышимыми искажениями.

Если фильтр LPC подключен до модуля преобразования MDCT, отбеленный сигнал преобразуется в MDCT-область. Поскольку этот сигнал имеет белый спектр, он не очень хорошо подходит для получения из него кривой перцепционного маскирования. Таким образом, кривая выравнивания усиления в MDCT-области, созданная для компенсации отбеливания спектра, может использоваться при определении пороговой кривой маскирования и/или масштабных коэффициентов. По этой причине масштабные коэффициенты должны определяться для сигнала, имеющего свойства абсолютного спектра исходного сигнала, чтобы правильно определить перцепционное маскирование. Вычисление кривой выравнивания усиления для MDCT-области из полинома LPC обсуждается ниже более подробно со ссылкой на фиг. 10.

Вариант осуществления приведенной выше в общих чертах схемы определения

масштабных коэффициентов представлен на фиг. 9а. В этом варианте осуществления входной сигнал вводится в модуль 901 LP, который определяет огибающую спектра входного сигнала, описанного с помощью $A(z)$, и выводит упомянутый полином, а также отфильтрованную версию входного сигнала. Входной сигнал фильтруется с помощью инверсии $A(z)$, чтобы обеспечить спектрально белый сигнал, который впоследствии используется другими частями кодера. Отфильтрованный сигнал $\hat{x}(n)$ вводится в блок 902 преобразования MDCT, тогда как полином $A(z)$ вводится в блок 970 вычисления кривой усиления MDCT 970 (как показано на фиг. 14). Кривая усиления, определенная из полинома LP, применяется к коэффициентам MDCT или линиям, чтобы сохранить спектральную огибающую первоначального входного сигнала до того, как определять масштабные коэффициенты. Линии MDCT с отрегулированным усилением вводятся в модуль 960 определения масштабных коэффициентов, который определяет масштабные коэффициенты для входного сигнала.

Используя кратко представленный выше подход, данные, передаваемые между кодером и декодером, содержит как полином LP, из которого может быть получена соответствующая перцепционная информация, так и модель сигнала, которая может быть получена, когда используется устройство квантования, основанное на модели, и масштабные коэффициенты обычно используются в трансформном кодеке.

Более подробно, возвращаясь к фиг. 9, блок 901 LPC, показанный на чертеже, определяет из входного сигнала огибающую спектра $A(z)$ сигнала и получает из нее перцепционное представление $A'(z)$. Кроме того, масштабные коэффициенты, которые обычно используются в основанных на преобразовании перцепционных аудиокодеках, определяются по входному сигналу или могут быть определены на белом сигнале, созданном фильтром LP, если передаточная функция фильтра LP учитывается при определении масштабных коэффициентов (как описано ниже в контексте фиг. 10). Масштабные коэффициенты могут затем адаптироваться в модуле 961 адаптации масштабных коэффициентов для данного полинома LP, как будет описано ниже, чтобы уменьшить скорость передачи данных, требующуюся для передачи масштабных коэффициентов.

Обычно масштабные коэффициенты передаются на декодер и, таким образом, появляется полином LP. Теперь, при условии, что и то, и другое определено из первоначального входного сигнала и что и то, и другое в какой-то степени коррелированы со свойствами абсолютного спектра первоначального входного сигнала, предлагается кодировать дельта-представление между ними, чтобы удалить любую избыточность, которая может возникнуть, если то и другое передаются отдельно. В соответствии с вариантом осуществления, эта корреляция используется следующим образом. Поскольку полином LPC, когда правильно подвергается линейной частотной модуляции и наклоняется, стремится представить пороговую кривую маскирования, два представления могут быть объединены, так чтобы переданные масштабные коэффициенты трансформного кодера представляли разность между желательными масштабными коэффициентами и теми, которые могут быть получены из переданного полинома LPC. Модуль 961 адаптации масштабных коэффициентов, показанный на фиг. 9, поэтому вычисляет разность между желательными масштабными коэффициентами, созданными из первоначального входного сигнала, и масштабными коэффициентами, полученным из LPC. Этот аспект сохраняет способность иметь устройство квантования на MDCT-основе, имеющее представление масштабных коэффициентов, которое обычно используется в трансформных кодерах в пределах структуры LPC, работающих на остатке LPC, и все еще имеет возможность переключения

на модели устройство квантования, основанное на модели, которое получает размеры шага квантования исключительно из данных линейного предсказания.

На фиг. 9b приведена упрощенная блок-схема кодера и декодера, соответствующая варианту осуществления. Входной сигнал в кодере пропускается через модуль 901 LPC, создающий отбеленный остаточный сигнал и соответствующие параметры линейного предсказания. Дополнительно, в модуле 901 LPC может содержаться нормализация усиления. Остаточный сигнал от LPC преобразуется в частотную область с помощью преобразования 902 MDCT. В правой части фиг. 9b изображен декодер. Декодер принимает линии квантованного MDCT, деквантует 911 их и применяет инверсное преобразование 912 MDCT с последующим синтезом LPC с помощью фильтра 913.

Отбеленный сигнал в качестве выходного сигнала модуля 901 LPC в кодере на фиг. 9b подается на вход блока 902 фильтров MDCT. Линии MDCT, как результат анализа MDCT, трансформно кодируются с помощью алгоритмом трансформного кодирования, состоящего из перцепционной модели, которая управляет желательным размером шага квантования для различных частей спектра MDCT. Значения, определяющие размер шага квантования, называются масштабными коэффициентами и существует одно значение масштабного коэффициента, необходимое для каждого элемента спектра MDCT, называемое полосой масштабного коэффициента. На предшествующем уровне техники алгоритмов трансформного кодирования масштабные коэффициенты передаются через битовый поток на кодер.

В соответствии с одним из аспектов изобретения, кривая перцепционного маскирования, определяемая из параметров LPC, как объяснялось со ссылкой на фиг. 9, применяется при кодировании масштабных коэффициентов, используемых в квантовании. Другая возможность определения кривой перцепционного маскирования состоит в использовании немодифицированных коэффициентов фильтра LPC для определения распределения энергии по линиям MDCT. Имея такую оценку энергии, психоакустическая модель, используемая в схемах трансформного кодирования, может быть применена как в кодере, так и в декодере, для определения кривой маскирования.

Два представления кривой маскирования затем объединяются, так что масштабные коэффициенты, которые должны быть переданы трансформным кодером, представляют разность между желательными масштабными коэффициентами и масштабными коэффициентами, которые могут быть получены из переданного полинома LPC или быть основаны на психоакустической модели LPC. Этот признак сохраняет способность иметь устройство квантования, основанное на MDCT, имеющее такое представление масштабных коэффициентов, которое обычно используется в трансформных кодерах, внутри структуры LPC, работающей с LPC-остатком, и все еще иметь возможность управлять шумом квантования на основе полосы масштабного коэффициента в соответствии с психоакустической моделью трансформного кодера. Преимущество состоит в том, что передача разности масштабных коэффициентов будет стоить меньше битов по сравнению с передачей абсолютных значений масштабных коэффициентов, не принимая во внимание уже существующие данные LPC. В зависимости от скорости передачи данных, размера кадра или других параметров, может быть выбрана величина остатка масштабных коэффициентов, который будет передаваться. Чтобы иметь полный контроль над полосой каждого масштабного коэффициента, дельта-параметр масштабного коэффициента может передаваться с помощью соответствующей схемы бесшумового кодирования. В других случаях стоимость передачи масштабных коэффициентов может быть дополнительно снижена за счет более грубого представления разностей масштабных коэффициентов. Особым случаем с самыми низкими накладными

затратами является случай, когда разность масштабных коэффициентов устанавливается на 0 для всех полос и никакая дополнительная информация не передается.

На фиг. 10 показан предпочтительный вариант осуществления перевода полиномов LPC в кривую усиления MDCT. Как показано на фиг. 2, MDCT работает с отбеленным сигналом, причем отбеливание выполняется фильтром 1001 LPC. Чтобы сохранить огибающую спектра первоначального входного сигнала, кривая усиления MDCT вычисляется модулем 1070 кривой усиления MDCT. Кривая выравнивания усиления в MDCT-области может быть получена посредством определения величины отклика огибающей спектра, описанной фильтром LPC, для частот, представленных элементами при MDCT-преобразовании. Кривая усиления может затем быть применена к данным MDCT, например, при вычислении минимального сигнала среднеквадратичной ошибки, как показано на фиг. 3, или при определении кривой перцепционного маскирования для определения масштабных коэффициентов, как показано выше со ссылкой на фиг. 9.

На фиг. 12 показан предпочтительный вариант осуществления адаптации вычисления для перцепционного взвешивающего фильтра, основанной на размере преобразования и/или типе устройства квантования. Полином LP $A(z)$ определяется модулем 1201 LPC, показанным на фиг. 16. Модуль 1271 изменения параметров LPC принимает параметры LPC, такие как полином LPC $A(z)$, и создает перцепционный взвешивающий фильтр $A'(z)$, изменяя параметры LPC. Например, ширина полосы полинома LPC $A(z)$ расширяется и/или полином наклоняется. Входные параметры для модуля 1272 адаптации линейной частотной модуляции и наклона являются заданными по умолчанию значениями линейной частотной модуляции и наклона, ρ и γ . Они изменяются по заданным, заранее определенным правилам, основываясь на используемом размере преобразования и/или на используемой стратегии Q квантования. Измененные параметры линейной частотной модуляции и наклона ρ' и γ' являются входными данными модуля 1271 модификации параметров LPC, переводящего огибающую спектра входного сигнала, представленную с помощью $A(z)$, в кривую перцепционного маскирования, представленную с помощью $A'(z)$.

Далее будет объяснена стратегия квантования, обусловленного размером кадра, и квантования, основанного на модели, обусловленного различными параметрами в соответствии с вариантом осуществления изобретения. Одним из аспектов настоящего изобретения является то, что оно использует различные стратегии квантования для различных размеров преобразования или размеров кадров. Это показано на фиг. 13, где размер кадра используется как параметр выбора для использования устройства квантования, основанного на модели, или устройства квантования, не основанного на модели. Следует отметить, что этот аспект квантования независим от других аспектов раскрытого кодера/декодера и может применяться также и в других кодеках. Примером устройства квантования, не основанного на модели, является устройство квантования на основе таблицы Хаффмана, используемое в стандарте аудиокодирования AAC. Устройство квантования, основанное на модели, может быть устройством квантования с ограничением энтропии (ECQ), использующим арифметическое кодирование. Однако, в вариантах осуществления настоящего изобретения могут также использоваться и другие устройства квантования.

В соответствии с независимым аспектом настоящего изобретения, предлагается переключение между различными стратегиями квантования как функция размера кадра, чтобы иметь возможность использовать оптимальную стратегию квантования, задаваемую конкретным размером кадра. Как пример, последовательность окон может

диктовать использование длительного преобразования для очень постоянного тонального музыкального сегмента сигнала. Для этого конкретного типа сигнала, использующего длительное преобразование, очень выгодно использовать стратегию квантования, с выгодой использующую символ "разреженности" (то есть, четко
 5 определенные дискретные тоны) в сигнальном спектре. Способ квантования, используемый в ААС, в комбинации с таблицами Хаффмана и группированием спектральных линий, также используемым в ААС, очень выгоден. Однако, с другой стороны, для речевых сегментов последовательность окон, учитывая выигрыш кодирования, даваемый ЛТР, может диктовать использование быстрых преобразований.
 10 Для этого типа сигнала и размера преобразования выгодно применять стратегию квантования, не пытающуюся находить или вводить разреженность в спектр, а вместо этого поддерживает энергию в широкой полосе, которая, учитывая ЛТР, будет сохранять подобный импульсу символ первоначального входного сигнала.

Более общее визуальное представление этой концепции приводится на фиг. 14, где
 15 входной сигнал преобразуется в MDCT-область и впоследствии квантуется устройством квантования, управляемым размером преобразования или размером кадра, используемым для преобразования MDCT.

В соответствии с другим аспектом изобретения, размер шага устройства квантования адаптируется как функция данных LPC и/или ЛТР. Это позволяет определять размер
 20 шага в зависимости от сложности кадра и управлять числом битов, выделяемых для кодирования кадра. На фиг. 15 показан пример того, как квантование, основанное на модели, может управляться данными LPC и ЛТР. В верхней части фиг. 15 приведена схематическая визуализация линий MDCT. Ниже показан размер шага квантования, дельта Δ , как функция частоты. Из этого конкретного примера ясно, что размер шага
 25 квантования увеличивается с частотой, то есть, для более высоких частот при квантовании вносится большее искажение. Дельта-кривая получается из параметров LPC и ЛТР посредством модуля дельта-адаптации, показанного на фиг. 15а. Дельта-кривая может дополнительно быть получена из полинома $A(z)$ предсказания, посредством линейной частотной модуляции и/или наклона, как было объяснено со
 30 ссылкой на фиг. 13.

Предпочтительная функция перцепционного взвешивания, полученная из данных LPC, определяется следующим уравнением:

$$35 \quad P(z) = \frac{1 - (1 - \tau)r_1 z^{-1}}{A(z/\rho)},$$

где $A(z)$ — полином LPC, τ - параметр наклона, ρ - управляет линейной частотной модуляцией, и r_1 - первый коэффициент отражения, вычисленный из полинома $A(z)$.

Следует отметить, что полином $A(z)$ может вычисляться повторно для подбора различных представлений, чтобы извлечь из полинома соответствующую информацию.
 40 Если Вы заинтересованы в крутизне спектра, чтобы применить "наклон" противодействия крутизне спектра, предпочтительно повторное вычисление полинома в коэффициенты отражения, так как первый коэффициент отражения представляет крутизну спектра.

Кроме того, дельта-значения Δ могут адаптироваться как функция дисперсии
 45 входного сигнала Δ , усиления ЛТР g и первого коэффициента отражения r_1 , полученного из полинома предсказания. Например, адаптация может основываться на следующем уравнении:

$$\Delta' = \Delta(1+r_1(1-g^2))$$

Ниже описываются аспекты устройств квантования, основанные на модели, соответствующие варианту осуществления настоящего изобретения. На фиг. 16 представлен один из аспектов устройства квантования, основанного на модели. Линии MDCT являются входным сигналом устройства квантования, использующего однородные скалярные устройства квантования. Кроме того, в устройство квантования вводятся случайные сдвиги, используемые как значения сдвига для интервалов квантования, сдвигающие границы интервалов. Предложенное устройство квантования обеспечивает преимущества векторного квантования, сохраняя способность скалярных устройств квантования к поиску. Устройство квантования выполняет итерации по ряду различных значений сдвига и вычисляет для них ошибку квантования. Значение сдвига (или вектор значения сдвига), минимизирующее искажение квантования для конкретных квантуемых линий MDCT, используется для квантования. Значение сдвига затем передается на декодер наряду с квантованными линиями MDCT. Использование случайных сдвигов вносит шумовое заполнение в деквантованный декодированный сигнал и, тем самым, позволяет избежать спектральных провалов в квантованном спектре. Это особенно важно для низких скоростей передачи данных, когда многочисленные линии MDCT иначе квантуются в нулевое значение, которое может приводить к слышимым провалам в спектре реконструированного сигнала.

На фиг. 17 схематично показано устройство квантования линий MDCT, основанное на модели (MBMLQ), соответствующее варианту осуществления изобретения. В верхней части фиг. 17 показан кодер 1700 MBMLQ. Кодер 1700 MBMLQ получает в качестве входного сигнала линии MDCT в кадре MDCT или линии MDCT остаточного LTP, если LTP присутствует в системе. MBMLQ использует статистические модели линий MDCT и исходные коды адаптируются к свойствам сигнала на покадровой основе MDCT, давая в результате эффективное сжатие битового потока.

Локальное усиление линий MDCT может определяться как среднеквадратичное значение линий MDCT и линий MDCT, нормализованных в модуле 1720 нормализации усиления перед вводом в кодер 1700 MBMLQ. Локальное усиление нормализует линии MDCT и является дополнением к нормализации усиления LP. Поскольку усиление LP адаптируется к изменениям уровня сигнала на большей временной шкале, локальное усиление адаптируется к изменениям на меньшей временной шкале, давая в результате улучшенное качество переходных звуков и начал в речи. Локальное усиление кодируется с фиксированной частотой или переменной частотой кодирования и передается на декодер.

Модуль 1710 управления частотой может использоваться для управления числом битов, используемых для кодирования кадра MDCT. Индекс управления частотой управляет числом используемых битов. Индекс управления частотой указывается в списке номинальных размеров шагов устройства квантования. Таблица может быть сортирована по размерам шага в порядке убывания (смотрите фиг. 17g).

Кодер MBMLQ работает с набором различных индексов управления частотой и для кадра используется индекс управления частотой, дающий в результате число битов, меньшее, чем число битов, предоставленных управлением битовым накопителем. Индекс управления частотой медленно изменяется и это может использоваться для уменьшения сложности поиска и эффективного кодирования индекса. Набор индексов, который проверен, может быть уменьшен, если проверка начинается вблизи индекса предыдущего кадра MDCT. Аналогично, эффективное энтропийное кодирование индекса получается, если вероятности достигают максимума вокруг предыдущего значения индекса.

Например, для списка из 32 размеров шага индекс управления частотой может быть кодирован, используя, в среднем, 2 бита на кадр MDCT.

На фиг. 17 дополнительно схематично показан декодер 1750 MBMLQ, где кадр MDCT повторно нормализуется по усилению, если в кодере 1700 было определено локальное усиление.

На фиг. 17а более подробно схематично показан кодер 1700 линий MDCT, основанный на модели, соответствующий варианту осуществления. Он содержит модуль 1730 предварительной обработки устройства квантования (смотрите фиг. 17с), кодер 1740, основанный на модели, с ограничением энтропии (смотрите фиг. 17е) и арифметический кодер 1720, который может быть арифметическим кодером, соответствующим предшествующему уровню техники. Задача модуля 1730 предварительной обработки устройства квантования состоит в том, чтобы адаптировать кодер MBMLQ к сигнальной статистике на покадровой основе MDCT. Он принимает в качестве входного сигнала другие параметры кодека и извлекает из них полезную статистику о сигнале, которая может использоваться для изменения кодера 1740, основанного на модели, с ограниченной энтропией. Кодер 1740, основанный на модели, с ограничением энтропии управляется, например, набором параметров управления: размер Δ шага устройства квантования, набор V оценок дисперсии строк MDCT (вектор; одно предполагаемое значение на линию MDCT), кривая перцепционного маскирования, P_{mod} , матрица или таблица (случайных) сдвигов и статистическая модель строк MDCT, которые описывают форму распределения строк MDCT и их взаимозависимости. Все упомянутые выше параметры управления могут изменяться между кадрами MDCT.

На фиг. 17b схематично показан декодер 1750 линий MDCT, основанный на модели, соответствующий варианту осуществления изобретения. В качестве входного сигнала он принимает побочные информационные биты из битового потока и декодирует их в параметры, являющиеся входными для модуля 1760 предварительной обработки устройства квантования (смотрите фиг. 17с). Модуль 1760 предварительной обработки устройства квантования предпочтительно имеет те же самые функциональные возможности в кодере 1700, как в декодере 1750. Параметры, которые вводятся в модуль 1760 предварительной обработки устройства квантования, являются точно такими же как в кодере, как и в декодере. Модуль 1760 предварительной обработки устройства квантования имеет на выходе набор параметров управления (таких же, как в кодере 1700) и они являются входным сигналом модуля 1770 вычислений вероятности (смотрите фиг. 17g; таким же, как в кодере, смотрите фиг. 17е) и модуля 1780 деквантования (смотрите фиг. 17h; то же самое в кодере, смотрите фиг. 17е). Таблицы cdf из модуля 1770 вычислений вероятности, представляющие плотности распределения вероятности для всех линий MDCT при заданных дельта-параметрах, используемых для квантования, и дисперсии сигнала, являются входными данными на арифметический декодер (который может быть любым арифметическим кодером, известным специалистам в этой области техники), который затем декодирует биты линий MDCT в индексы линий MDCT. Индексы линий MDCT затем деквантуются в линии MDCT модулем 1780 деквантования.

На фиг. 17с схематично показаны аспекты предварительной обработки устройства квантования в соответствии с вариантом осуществления изобретения, которая состоит из i) вычисления размера шага, ii) изменения кривой перцепционного маскирования, iii) определения дисперсии линий MDCT, iv) построения таблицы сдвигов.

Вычисление размера шага объясняется более подробно на фиг. 17d. Оно содержит i) поиск таблицы, где индекс управления частотой указывает в таблице размеров шага те размеры, которые создают номинальный Δ_{nom} , ii) адаптацию к низкой энергии, и iii)

высокочастотную адаптацию.

Нормализация усиления обычно приводит к тому, что звуки с высокой энергией и звуки с низкой энергией кодируются с одним и тем же сегментным отношением "сигнал/шум" (SNR). Это может привести к чрезмерному числу битов, используемых для звуков с низкой энергией. Предложенная адаптация к низкой энергии позволяет получить точную настройку компромисса между звуками с низкой энергией и высокой энергией. Размер шага может быть увеличен, когда энергия сигнала становится низкой, как показано на фиг. 17d-ii), где приводится пример кривой отношения между энергией сигнала (усиление g) и коэффициентом управления q_{Lc} . Усиление сигнала g может быть вычислено как среднеквадратичное значение самого входного сигнала или остаток LP. Кривая управления на фиг. 17d-ii) является только одним примером и для увеличения размера шага для сигналов с низкой энергией могут использоваться и другие функции управления. В показанном на чертеже примере функция управления определяется пошаговыми линейными участками, которые определяются порогами T_1 и T_2 , и коэффициентом размера шага L .

Высокочастотные звуки перцепционно менее важны, чем низкочастотные звуки. Функция высокочастотной адаптации увеличивает размер шага, когда кадр MDCT является высокочастотным, то есть, когда энергия сигнала в текущем кадре MDCT сконцентрирована на более высоких частотах, приводя в результате к меньшему числу битов, затраченных на такие кадры. Если LTP присутствует и если усиление LTP g_{LTP} близко к 1, остаток LTP может становиться высокочастотным; в таком случае выгодно не увеличивать размер шага. Этот механизм показан на фиг. 17d-iii), где r - первый коэффициент отражения, полученный из LPC. Предложенная высокочастотная адаптация может использовать следующее уравнение:

$$q_{hp} = \begin{cases} 1 + r(1 - g^2) & \text{если } r > 0 \\ 1 & \text{если } r \leq 0 \end{cases}$$

На фиг. 17c-ii) схематично показана модификация кривой перцепционного маскирования, использующей подъем усиления на низкой частоте (LF), чтобы удалить "громopodobные" артефакты кодирования. Увеличение усиления на низких частотах может быть фиксированным или делаться адаптивным, так чтобы дополнительное усиление получала только часть ниже первого спектрального пика. Увеличение усиления на низких частотах может адаптироваться, использовании данных огибающей LPC.

На фиг. 17c-iii) схематично показано определение дисперсии линий MDCT. При активном отбеливающем фильтре LPC все линии MDCT имеют единичную дисперсию (соответственно огибающей LPC). После перцепционного взвешивания в кодере 1740, основанном на модели, с ограниченной энтропией (смотрите фиг. 17e) линии MDCT имеют дисперсии, являющиеся инверсией квадратичной кривой перцепционного маскирования или квадратичной модифицированной кривой маскирования P_{mod} . Если LTP присутствует, оно может уменьшить дисперсию линий MDCT. На фиг. 17c-iii) изображен механизм, который адаптирует определенные дисперсии к LTP. На чертеже показана функция модификации q_{LTP} по частоте f . Модифицированные дисперсии могут быть определены с помощью $V_{LTPmod} = V * q_{LTP}$. Значение L_{LTP} может быть функцией усиления LTP, так чтобы L_{LTP} был ближе к 0, если усиление LTP приблизительно равно 1 (указатель того, что LTP обнаружило хорошее соответствие), и L_{LTP} ближе к 1, если усиление LTP приблизительно равно 0. Предложенная адаптация LTP для дисперсий V

$= \{v_1, v_2, \dots, v_j, \dots, v_N\}$ влияет только на линии MDCT ниже определенной частоты ($f_{LTPcutoff}$). В результате, дисперсии линий MDCT ниже частоты среза $f_{LTPcutoff}$ уменьшаются, причем уменьшение зависит от усиления LTP.

5 На фиг. 17c-iv) схематично показана конструкция таблицы сдвигов. Номинальная таблица сдвигов является матрицей, заполненной псевдослучайными числами, распределенными между -0,5 и 0,5. Число столбцов в матрице равняется числу линий MDCT, которые кодируются с помощью MBMLQ. Число строк регулируется и равняется числу векторов сдвигов, которые проверяются при RD-оптимизации в кодере 1740, основанном на модели, с ограниченной энтропией (смотрите фиг. 17e). Функция
10 конструкции таблицы сдвигов масштабирует номинальную таблицу сдвигов с размером шага устройства квантования, так чтобы сдвиги были распределены между $-\Delta/2$ и $+\Delta/2$.

На фиг. 17g схематично показан вариант осуществления таблицы сдвигов. Индекс сдвига является указателем в таблице и выбирает нужный вектор сдвига $O = \{o_1, o_2, \dots, o_n, \dots, o_N\}$, где N - число линий MDCT во кадре MDCT.
15

Как описано ниже, сдвиги обеспечивают средство для шумового заполнения. Лучшее объективное и перцепционное качество получается, если разброс сдвигов ограничивается для линий MDCT, имеющих низкую дисперсию v_j по сравнению с размером Δ шага
20 устройства квантования. Пример такого ограничения описан на фиг. 17c-iv), где k_1 и k_2 являются настраиваемыми параметрами. Распределение сдвигов может быть равномерным и находиться между $-s$ и $+s$. Границы s могут определяться в соответствии с выражением:

$$25 \quad s = \begin{cases} k_2 \sqrt{v_j} & \text{если } \sqrt{v_j} < k_1 \Delta \\ \frac{\Delta}{2} & \text{в противном случае} \end{cases}$$

Для линий MDCT с низкой дисперсией (где v_j мало по сравнению с Δ) может быть
30 предпочтительным сделать распределение сдвигов неравномерным и зависимым от сигнала.

На фиг. 17e более подробно схематично показан кодер 1740, основанный на модели, с ограничением энтропии 1740. Входные линии MDCT перцепционно взвешены посредством деления их на значения кривой перцепционного маскирования,
35 предпочтительно полученной из полинома LPC, что приводит к взвешенным векторам линий MDCT $y = (y_1, \dots, y_N)$. Цель последующего кодирования состоит в том, чтобы ввести белый шум квантования в линии MDCT в перцепционной области. В декодере применена инверсия перцепционного взвешивания, который приводит к шуму квантования, соответствующему кривой перцепционного маскирования.

40 Сначала описывается итерация по случайным сдвигам. В матрице сдвигов для каждой строки j в матрице сдвигов выполняются следующие операции: Каждая строка MDCT квантуется однородными скалярными устройствами квантования (USQ) сдвигов, в которых каждое устройство квантования сдвигается на свое собственное уникальное значение сдвига, взятое из вектора строки сдвигов.

45 Вероятность минимального интервала искажения от каждого USQ вычисляется в модуле 1770 вычисления вероятности (смотрите фиг. 17g). Индексы USQ являются энтропийно кодированными. Стоимость с точки зрения числа битов, требующихся для кодирования индексов, вычисляется соответственно фиг. 17e, давая в результате длину

кодового слова R_j . Граница перегрузки USQ_j для линии MDCT может быть вычислена как $k_3 \sqrt{V_j}$, где k_3 может выбираться, чтобы быть любым соответствующим числом, например 20. Границей перегрузки является граница, на которой ошибка квантования по величине больше, чем половина размера шага квантования.

Скалярное значение реконструкции для каждой линии MDCT вычисляется модулем 1780 деквантования (смотрите фиг. 17h), давая в результате квантованный вектор y MDCT. В модуле 1790 оптимизации RD вычисляется искажение $D_j = d(y, \bar{y})$. $d(y, \bar{y})$ может быть среднеквадратичной ошибкой (MSE) или другой перцепционно более подходящей мерой искажения, например, основанной на перцепционной функции взвешивания. В частности, может быть полезна мера искажения, которая взвешивает вместе MSE и рассогласование по энергии между y и \bar{y} .

В модуле 1790 RD-оптимизации вычисляется стоимость C , предпочтительно основанная на искажении D_j и/или теоретической длине кодового слова R_j для каждой строки j в матрице сдвигов. Пример функции стоимости является $C = 10 * \log_{10}(D_j) + \lambda * R_j / N$. Сдвиг, которое минимизирует C , выбирается и из кодера 1780, основанного на модели, с ограниченной энтропией выводятся соответствующие индексы USQ и вероятности.

RD-оптимизация может быть по желанию дополнительно улучшена, изменяя другие свойства устройства квантования вместе со сдвигом. Например, вместо использования одного и того же фиксированного V оценки дисперсии для каждого вектора сдвига, которая проверяется при RD-оптимизации, вектор V оценки дисперсии может быть переменным. Для вектора m строки сдвигов может затем использоваться оценка дисперсии $k_m * V$, где k_m может охватывать, например, диапазон от 0,5 до 1,5 по мере того, как m изменяется от $m=1$ до $m=(\text{число строк в матрице сдвигов})$. Это делает энтропийное кодирование и вычисление MMSE менее чувствительными к изменениям в статистике входного сигнала, которые не может зафиксировать статистическая модель. Это приводит к более низкой стоимости C в целом.

Деквантованные линии MDCT могут быть дополнительно улучшены при использовании устройства квантования остатка, как показано на фиг. 17e. Устройство квантования остатка может быть, например, устройством квантования случайного вектора с фиксированной частотой.

Работа однородного скалярного устройства квантования (USQ) для квантования линии n MDCT схематично показана на фиг. 17f, где показано значение n линии MDCT, находящееся в интервале минимальных искажений, имеющем индекс i_n . Отметки "x" указывают центр (средняя точка) интервалов квантования с размером шага Δ . Начало координат скалярного устройства квантования сдвинуто на величину сдвига o_n от вектора сдвига $O = \{o_1, o_2, \dots, o_n, \dots, o_N\}$. Таким образом, границы интервала и средние точки сдвигаются на величину сдвига.

Использование сдвигов вносит в квантованный сигнал управляемое кодером шумовое заполнение и это позволяет избежать провалов в квантованном спектре. Дополнительно, сдвиги повышают эффективность кодирования, обеспечивая ряд альтернатив кодирования, которые заполняют пространство более эффективно, чем кубическая решетка. Кроме того, сдвиги обеспечивают изменение в таблицах вероятности, которые вычисляются модулем 1770 вычисления вероятности, ведущее к более эффективному энтропийному кодированию индексов линий MDCT (то есть, требуется меньше битов).

Использование переменного размера шага Δ (дельта) позволяет иметь переменную

точность при квантовании, так, чтобы более высокая точность могла использоваться для перцепционно важных звуков, и меньшая точность могла использоваться для менее важных звуков.

На фиг. 17g схематично показано вычисление вероятности в модуле 1770 вычисления вероятности. Входными сигналами этого модуля являются статистическая модель, применяемая для линий MDCT, размер шага Δ , вектор V дисперсии, индекс сдвига и таблица сдвигов. Выходом модуля 1770 вычисления вероятности являются таблицы cdf. Для каждой линии x_j MDCT определяется статистическая модель (то есть, функция плотности распределения вероятности, pdf). Область под функцией pdf для интервала I является вероятностью p_{ij} интервала. Эта вероятность используется для арифметического кодирования линий MDCT.

На фиг. 17h схематично показано, как выполняется процесс деквантования, например, в модуле 1780 деквантования. Центр масс (значение MMSE) x_{MMSE} для минимального интервала искажения каждой линии MDCT вычисляется вместе со средней точкой x_{MP} интервала. Полагая, что квантуется N -мерный вектор линий MDCT, скалярное значение MMSE является близким к оптимальному и, в целом, слишком низким. Это приводит к потере дисперсии и спектральному разбалансу в декодированном выходном сигнале. Эта проблема может быть смягчена декодированием с сохранением дисперсии, как описано на фиг. 17h, где значение реконструкции вычисляется как взвешенная сумма значения MMSE и значения в средней точке. Дополнительное необязательное улучшение состоит в адаптации веса так, чтобы значение MMSE доминировало для речи, а средняя точка доминировала для неречевых звуков. Это дает в результате более чистую речь, при том, что для неречевых звуков сохраняются спектральный баланс и энергия.

Декодирование с сохранением дисперсии в соответствии с вариантом осуществления изобретения достигается, определяя точку реконструкции в соответствии со следующим уравнением:

$$x_{dequant} = (1 - \Delta)x_{MMSE} + x_{MP}$$

Адаптивное декодирование с сохранением дисперсии может быть основано на следующем правиле определения коэффициента интерполяции:

$$x = \begin{cases} 0 & \text{для речевых звуков} \\ 1 & \text{для неречевых звуков} \end{cases}$$

Адаптивный вес дополнительно может быть функцией, например, усиления g_{LTP} при предсказании LTP: $\Delta = f(g_{LTP})$. Адаптивный вес медленно изменяется и может эффективно кодироваться рекурсивным энтропийным кодом.

Статистическая модель линий MDCT, используемая при вычислениях вероятности (фиг. 17g) и при деквантовании (фиг. 17h) должна отражать статистику реального сигнала. В одной версии статистическая модель предполагает, что линии MDCT независимы и имеют лапласово распределение. Другая версия моделирует линии MDCT как независимые гауссовы распределения. Одна версия моделирует линии MDCT как модели с гауссовой смесью, содержащие взаимозависимости между линиями MDCT внутри и между кадрами MDCT. Другая версия адаптирует статистическую модель к текущей статистике сигнала. Адаптивные статистические модели могут адаптироваться вперед и/или назад.

Другой аспект изобретения, относящийся к модифицированным точкам реконструкции устройства квантования, схематично показан на фиг. 19, где представлено инверсное устройство квантования, используемое в декодере варианта осуществления.

Модуль имеет, помимо обычных входных сигналов инверсного устройства квантования, то есть, квантованных линий и информации о размере шага квантования (типе квантования), также информацию о точке реконструкции устройства квантования.

Инверсное устройство квантования этого варианта осуществления может использовать
5 многочисленные типы точек реконструкции, когда реконструированное значение \bar{y}_n определяется из соответствующего индекса i_n квантования. Как упомянуто выше,

значения \bar{y}_n реконструкции дополнительно используются, например, в кодере линий

10 MDCT (смотрите фиг. 17), чтобы определить остаток квантования для ввода в устройство квантования остатка. Дополнительно, реконструкция квантования выполняется в инверсном устройстве 304 квантования при реконструкции кодированного кадра MDCT для использования в буфере LTP (смотрите фиг. 3) и, естественно, в декодере.

Инверсное устройство квантования может выбрать, например, среднюю точку
15 интервала квантования в качестве точки реконструкции или точки реконструкции MMSE.

В варианте осуществления настоящего изобретения точка реконструкции устройства квантования выбирается так, чтобы быть средним значением между центральной точкой и точкой реконструкции MMSE. В целом, точка реконструкции может быть интерполирована между средней точкой и точкой реконструкции MMSE,
20 например, в зависимости от свойств сигнала, таких как периодичность сигнала.

Информация о периодичности сигнала может быть получена, например, от модуля LTP. Этот признак позволяет системе управлять искажением и сохранением энергии.

Центральная точка реконструкции гарантирует сохранение энергии, тогда как точка реконструкции MMSE гарантирует минимальное искажение. С учетом сигнала, система
25 может затем адаптировать точку реконструкции туда, где обеспечивается наилучший компромисс.

Настоящее изобретение дополнительно содержит новый формат кода последовательности окон. В соответствии с вариантом осуществления изобретения, окна, используемые для преобразования MDCT, имеют диадические размеры и могут
30 изменяться по размеру от окна к окну только с коэффициентом два. Диадическими размерами преобразования являются, например, выборки 64, 128..., 2048, соответствующие 4, 8..., 128 мс при частоте дискретизации 16 кГц. В целом, предлагаются окна с переменным размером, которые могут принимать многочисленные размеры окон между минимальным размером и максимальным размером окон. В

35 последовательности, размеры последовательных окон могут изменяться только с коэффициентом два, так чтобы формировались гладкие последовательности размеров окна без резких изменений. Последовательности окон, как они определяются вариантом осуществления, то есть, ограниченные диадическими размерами и имеющие возможность изменяться по размеру от окна к окну только с коэффициентом два, имеют несколько
40 преимуществ. Во-первых, не требуется никакое специальное стартовое или стоповое окно, то есть, окна с резкими краями. Это способствует хорошей разрешающей способности по времени/частоте. Во-вторых, последовательность окон становится очень эффективной для кодирования, то есть, чтобы сигнализировать на декодер, какая конкретная последовательность окон используется. Наконец, последовательность окон
45 всегда будет точно вписываться в структуру гиперкадра.

Структура гиперкадра полезна при работе кодера в реальной системе, где для того, чтобы иметь возможность запустить декодер, должны передаваться определенные параметры конфигурации декодера. Эти данные обычно запоминаются в поле заголовка

в битовом потоке, описывающем кодированный аудиосигнал. Чтобы минимизировать скорость передачи данных, заголовок не передается для каждого кадра кодированных данных, особенно в системе, предложенной в соответствии с настоящим изобретением, где размеры кадра MDCT могут изменяться от очень короткого до очень длинного.

5 Поэтому в соответствии с настоящим изобретением предложено группировать определенное количество кадров MDCT вместе в один гиперкадр, в котором данные заголовка передаются в начале гиперкадра. Гиперкадр обычно определяется как имеющий определенную продолжительность во времени. Поэтому необходимо позаботиться о том, чтобы изменения размеров кадров MDCT вписывались в
10 постоянную длительность, заранее заданную длительность гиперкадра. Упомянутая выше соответствующая изобретению последовательность окон обеспечивает, что выбранная последовательность окон всегда вписывается в структуру гиперкадра.

В соответствии с вариантом осуществления настоящего изобретения, задержка LTP и усиление LTP кодируются в виде переменной частоты. Это предпочтительно, поскольку
15 из-за эффективности LTP для постоянных периодических сигналов задержка LTP имеет тенденцию быть одной и той же для отчасти длинных сегментов. Следовательно, это может быть использовано средством арифметического кодирования, приводя в результате к кодированию задержки LTP и усиления LTP с переменной частотой.

Аналогично, вариант осуществления настоящего изобретения использует
20 преимущества битового накопителя и кодирования с переменной скоростью для кодирования параметров LP. Кроме того, настоящим изобретением представляется рекурсивное кодирование LP.

Другой аспект настоящего изобретения состоит в работе с битовым накопителем для переменных размеров кадра в кодере. На фиг. 18 представлен блок 1800 управления
25 битовым накопителем, соответствующий настоящему изобретению. В дополнение к мере сложности, предоставленной в качестве входного сигнала, блок управления битовым накопителем также принимает информацию о длительности текущего кадра. Примером меры сложности для использования в блоке управления битовым накопителем является перцепционная энтропия или логарифм энергетического спектра. Управление
30 битовым накопителем важно в системе, где длительности кадров могут меняться согласно набору различных длительностей кадров. Предложенный блок 1800 управления битовым накопителем учитывает длительность кадра при вычислении числа предоставляемых битов для кадра, который должен кодироваться, как будет описано ниже.

35 Битовый накопитель определяется здесь как определенное фиксированное количество битов в буфере, которое должно быть больше, чем среднее число битов, которые разрешается использовать кадру для заданной скорости передачи данных. Если он имеет тот же самый размер, то никакое изменение в числе битов для кадра не может быть возможно. Управление битовым накопителем всегда проверяет уровень битового
40 накопителя перед извлечением битов, которые будут предоставлены алгоритму кодирования в качестве разрешенного числа битов для фактического кадра. Таким образом, полный битовый накопитель означает, что число битов, доступных в битовом накопителе, равняется размеру битового накопителя. После кодирования кадра число использованных битов будет вычитаться из буфера и битовый накопитель будет
45 обновляться добавлением некоторого числа битов, представляющих постоянную битовую частоту. Поэтому битовый накопитель пуст, если число битов в битовом накопителе перед кодированием кадра равно среднему числу битов на кадр.

На фиг. 18a представлена базовая концепция управления битовым накопителем.

Кодер обеспечивает средство вычисления того, насколько сложно кодировать фактический кадр по сравнению с предыдущим кадром. Для средней сложности 1,0, число предоставленных битов зависит от числа битов, доступных в битовом накопителе. В соответствии с заданной линией управления, число, большее битов, чем число битов, соответствующее средней частоте передачи данных, будет изыматься из битового накопителя, если битовый накопитель будет совершенно полон. В случае пустого битового накопителя, для кодирования кадра будет использоваться меньшее число битов по сравнению со средним числом битов кадра. Такое поведение приводит в результате к среднему уровню битового накопителя для более длинной последовательности кадров со средней сложностью. Для кадров с более высокой сложностью линия управления может быть смещена вверх, приводя к эффекту, при котором при сложном кодировании кадров разрешается использовать больше битов при том же самом уровне битового накопителя. Соответственно, для облегчения кодирования кадров число битов, разрешенных для кадра, должно быть уменьшено простым сдвигом вниз на линию управления на фиг. 18а относительно случая средней сложности к случаю меньшей сложности. Возможны также другие модификации, помимо простого смещения линии управления. Например, как показано на фиг. 18а, крутизна кривой управления может изменяться в зависимости от сложности кадров.

При вычислении числа предоставленных битов необходимо подчиняться пределам по нижней границе битового накопителя, чтобы не взять из буфера больше битов, чем разрешено. Схема управления битовым накопителем, содержащая вычисление предоставленных битов с помощью линии управления, как показано на фиг. 18а, является только одним примером возможной меры уровня битового накопителя и сложности в соотношениях предоставленных битов. Другие алгоритмы управления будут также иметь, в целом, жесткие пределы на нижней границе уровня битового накопителя, которые не позволяют битовому накопителю нарушать ограничение по опустошению битового накопителя, а также пределы на верхней границе, где кодер будет принуждаться записывать заполняющие биты, если кодером будет расходоваться слишком малое число битов.

Для такого механизма управления, способного обрабатывать набор переменных размеров кадра, этот простой алгоритм управления должен быть адаптирован. Используемая мера сложности должна быть нормализована, так чтобы значения сложности различных размеров кадра были сопоставимы. Для каждого размера кадра будет существовать различный допустимый диапазон предоставляемых битов и поэтому среднее число битов на кадр различно для переменного размера кадра, следовательно, каждый размера кадра имеет свое собственное уравнение управления со своими собственными ограничениями. Один пример показан на фиг. 18b. Важной модификацией случая с фиксированным размером кадра является пониженная допустимая граница алгоритма управления. Вместо среднего числа битов для фактического размера кадра, которое соответствует установленному случаю фиксированной скорости передачи данных, теперь среднее число битов для наибольшего допустимого размера кадра является наинизшим допустимым значением для уровня битового накопителя перед извлечением битов для фактического кадра. Это одно из основных отличий для управления битовым накопителем для кадров с фиксированными размерами. Это ограничение гарантирует, что следующий кадр с наибольшим возможным размером кадра сможет использовать, по меньшей мере, среднее число битов для этого размера кадра.

Мера сложности может основываться, например, на вычислении перцепционной

энтропии (PE), которая получается из порогов маскирования психоакустической модели, как это делается в AAC, или, как альтернатива, из числа битов квантования с фиксированным размером шага, как это делается в части ECQ кодера в соответствии с вариантом осуществления настоящего изобретения. Эти значения могут быть
 5 нормализованы в отношении переменных размеров кадров, что может достигаться простым делением на длительность кадра и результатом будет соответствующее PE число битов на выборку. Другой этап нормализации может иметь место в отношении средней сложности. Для этой цели может использоваться скользящее среднее значение по прежним кадрам, приводящее в результате к значению сложности, большему, чем
 10 1,0, для сложных кадров или меньшему, чем 1,0, для простых кадров. В случае кодера с двумя проходами или большого просмотра вперед, значения сложности будущих кадров также могут быть учтены для этой нормализации меры сложности.

Другой аспект изобретения относится к специфическим особенностям работы с битовым накопителем для ECQ. Управление битовым накопителем для ECQ работает
 15 согласно предположению, что ECQ создает приблизительно постоянное качество, когда использует для кодирования постоянный размер шага устройства квантования. Постоянный размер шага устройства квантования создает переменную частоту и задача битового накопителя состоит в том, чтобы поддерживать как можно меньшим изменение в размере шага устройства квантования для различных кадров, не нарушая ограничения
 20 для буфера битового накопителя. В дополнение к частоте, созданной ECQ, на основе MDCT-кадра передается дополнительная информация (например, усиление и задержка LTP). Дополнительная информация, в целом, также энтропийно кодирована и, таким образом, использует различную частоту от кадра к кадру.

В варианте осуществления изобретения предложенное управление битовым
 25 накопителем пытается минимизировать изменение размера шага ECQ, вводя три переменные (смотрите фиг. 18c):

- R_{ECQ_AVG} : средняя частота ECQ на выборку, использованная ранее;
- Δ_{ECQ_AVG} : средний размер шага устройства квантования, использованный ранее.

Обе эти переменные динамически обновляются, чтобы отражать последнюю
 30 статистику кодирования.

- $R_{ECQ_AVG_DES}$: частота ECQ, соответствующая средней общей скорости передачи данных.

Это значение будет отличаться от R_{ECQ_AVG} в случае, если уровень битового
 35 накопителя изменился во время временного кадра окна усреднения, например, во время этого временного кадра использовалась скорость передачи данных, которая выше или ниже, чем указанная средняя скорость передачи данных. Оно также обновляется по мере того, как изменяется частота побочной информации, так чтобы общая частота равнялась указанной скорости передачи данных.

Управление битовым накопителем использует эти три значения, чтобы определить
 40 начальное предположение по дельте, которая должна использоваться для текущего кадра. Это делает путем нахождения $\Delta_{ECQ_AVG_DES}$ на кривой R_{ECQ_Δ} , показанной на фиг. 18c, которая соответствует $R_{ECQ_AVG_DES}$. На втором этапе это значение, возможно, изменяется, если частота не соответствует ограничениям битового накопителя. Пример
 45 кривой R_{ECQ_Δ} , показанной на фиг. 18C, основан на следующем уравнении:

Конечно, могут также использоваться другие математические соотношения между R_{ECQ} и Δ .

В стационарном случае, R_{ECQ_AVG} будет близко к $R_{ECQ_AVG_DES}$, и изменение в Δ будет очень маленьким. В нестационарном случае, операция усреднения гарантирует плавное изменение Δ .

5 Хотя вышесказанное было раскрыто в отношении конкретных вариантов осуществления настоящего изобретения, подразумевается, что концепция изобретения не ограничивается описанными вариантами осуществления. С другой стороны, раскрытие, представленное в настоящей заявке, даст возможность специалистам в этой области техники понять и выполнить изобретение. Специалистам в этой области техники должно быть понятно, что могут быть внесены различные изменения, не отступая от
10 сущности и объема изобретения, исключительно как они изложены в сопроводительной формуле изобретения.

(57) Формула изобретения

1. Аудиокодер, содержащий:

15 блок (201) линейного предсказания (LP) для анализа кадров первой длительности аудиосигнала для определения последовательности полиномов LPC;

блок (202) адаптивного преобразования длительности для преобразования кадров переменной второй длительности аудиосигнала в кадры коэффициентов модифицированного дискретного косинусного преобразования (MDCT);

20 блок (970, 1070) создания кривой усиления для создания кривых усиления MDCT-области для кадров коэффициентов MDCT на основе величин откликов, определенных из последовательности полиномов LPC, причем создание кривых усиления в MDCT-области содержит отображение, посредством блока (1100) отображения, полиномов LPC в соответствующие кадры коэффициентов MDCT;

25 блок применения кривой усиления для применения кривых усиления в MDCT-области к кадрам коэффициентов MDCT для создания кадров коэффициентов MDCT, скорректированных на усиление; и

блок (203) квантования для квантования кадров коэффициентов MDCT, скорректированных на усиление.

30 2. Аудиокодер по п. 1, содержащий

блок управления последовательностью окон для определения, для блока аудиосигнала, длительностей второго кадра для перекрытия окон MDCT.

3. Аудиокодер по любому из предыдущих пунктов, содержащий блок перцептуального моделирования, изменяющий характеристику фильтра LP посредством линейной частотной модуляции и/или наклона полинома LPC, созданного блоком линейного
35 предсказания для кадра LPC.

4. Аудиокодер по любому из предыдущих пунктов, содержащий:

блок разделения частот для разделения аудиосигнала на низкочастотный компонент и высокочастотный компонент; и

40 высокочастотный кодер для кодирования высокочастотного компонента, причем низкочастотный компонент является входным сигналом для блока линейного предсказания и блока преобразования, и причем высокочастотный кодер является кодером репликации спектральной полосы.

5. Аудиокодер по п. 4, в котором блок разделения частот содержит хранилище квадратурных зеркальных фильтров и блок синтеза квадратурных зеркальных фильтров, выполненные с возможностью понижающей дискретизации аудиосигнала.

6. Аудиокодер по п. 4 или 5, в котором граница между низкочастотной полосой и высокочастотной полосой может меняться и блок разделения частот определяет частоту

разделения на основе свойств аудиосигнала и/или требований к полосе пропускания кодера.

7. Аудиокодер по любому из предыдущих пунктов, содержащий блок (1360) определения масштабных коэффициентов для определения масштабных коэффициентов, чтобы управлять шумом квантования блока (203) квантования.

8. Аудиокодер по п. 7, в котором масштабные коэффициенты определяются на основе преобразованных кривых усиления в MDCT-области.

9. Аудиокодер по любому из предыдущих пунктов, содержащий параметрический стереоблок для вычисления параметрического стереопредставления левого и правого входных каналов.

10. Аудиокодер по любому из предыдущих пунктов, в котором блок (1500) отображения интерполирует полиномы LP, созданные на скорости, соответствующей длительности первого кадра, так чтобы они соответствовали кадрам сигнала в MDCT-области, созданного на скорости, соответствующей длительности второго кадра.

11. Аудиодекодер, содержащий:

блок (211) деквантования для деквантования квантованных кадров коэффициентов MDCT, декодированных из входного битового потока, причем кадры коэффициентов MDCT представляют аудиосигнал;

блок (970, 1070) создания кривой усиления для создания кривых усиления MDCT-области для кадров коэффициентов MDCT на основе величин откликов, определенных из последовательности полиномов LPC, причем полиномы LPC получены из входного битового потока, и причем создание кривых усиления MDCT-области содержит отображение, блоком (1100) отображения, полиномов LPC в соответствующие кадры коэффициентов MDCT;

блок применения кривой усиления для применения кривых усиления в MDCT-области к кадрам коэффициентов MDCT для создания кадров коэффициентов MDCT, скорректированных на усиление; и

блок (212) инверсного преобразования MDCT с адаптивной длиной для инверсного преобразования кадров коэффициентов MDCT, скорректированных на усиление в аудиосигнал временной области, причем блок инверсного MDCT-преобразования работает с переменной длительностью кадров.

12. Способ аудиокодирования, содержащий этапы, на которых:

выполняют анализ линейного предсказания (LP) для кадров первой длительности аудиосигнала для определения последовательности полиномов LPC;

преобразуют кадры переменной второй длительности аудиосигнала в кадры коэффициентов модифицированного дискретного косинусного преобразования (MDCT);

создают кривые усиления MDCT-области для кадров коэффициентов MDCT на основе величин откликов, определенных из последовательности полиномов LPC, причем создание кривых усиления в MDCT-области содержит отображение полиномов LPC в соответствующие кадры коэффициентов MDCT;

применяют кривые усиления в MDCT-области к кадрам коэффициентов MDCT для получения кадров коэффициентов MDCT, скорректированных на усиление; и

квантуют кадры коэффициентов MDCT, скорректированных на усиление.

13. Способ аудиодекодирования, содержащий этапы, на которых:

деквантуют квантованные кадры коэффициентов MDCT, декодированных из входного битового потока, причем кадры коэффициентов MDCT представляют аудиосигнал;

создают кривые усиления MDCT-области для кадров коэффициентов MDCT на основе величин откликов, определенных из полиномов LPC, причем полиномы LPC получены

из входного битового потока, и причем создание кривых усиления MDCT-области содержит отображение полиномов LPC в соответствующие кадры коэффициентов MDCT;

5 применяют кривые усиления в MDCT-области к кадрам коэффициентов MDCT для создания кадров коэффициентов MDCT, скорректированных на усиление;

инверсно преобразуют кадры коэффициентов MDCT, скорректированных на усиление, в аудиосигнал временной области с использованием инверсного MDCT, работающего с переменной длительностью кадров.

10 14. Программируемое устройство, содержащее компьютерную программу для побуждения программируемого устройства выполнять способ аудиокодирования по п.12.

15 15. Программируемое устройство, содержащее компьютерную программу для побуждения программируемого устройства выполнять способ аудиодекодирования по п.13.

16. Система аудиокодирования, содержащая аудиокодер по п.1 и аудиодекодер по п.11.

20

25

30

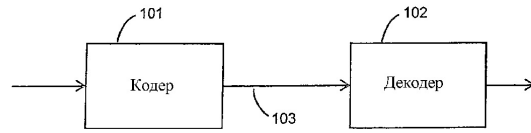
35

40

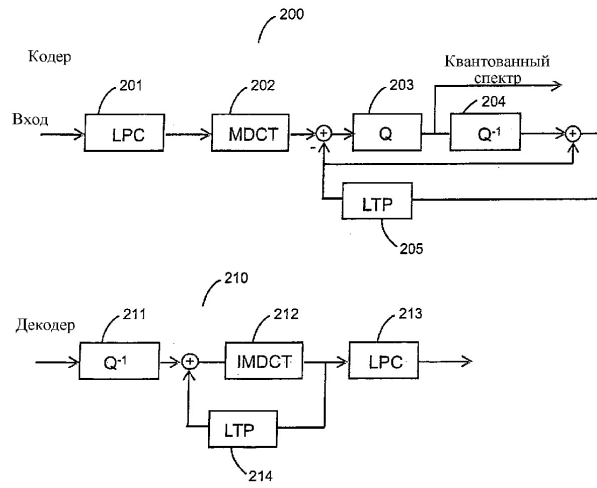
45

1

1/21

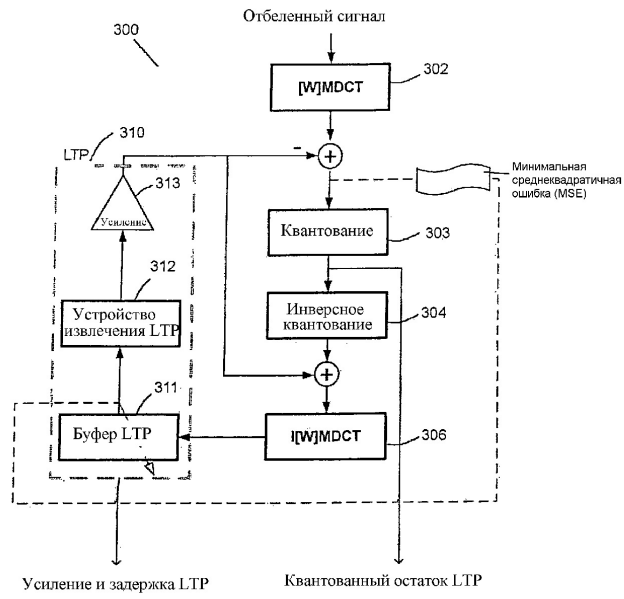


Фиг. 1

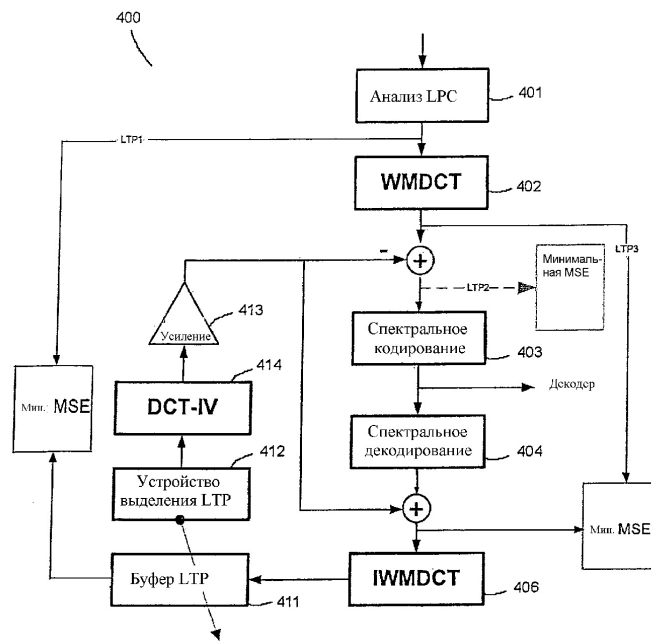


Фиг. 2

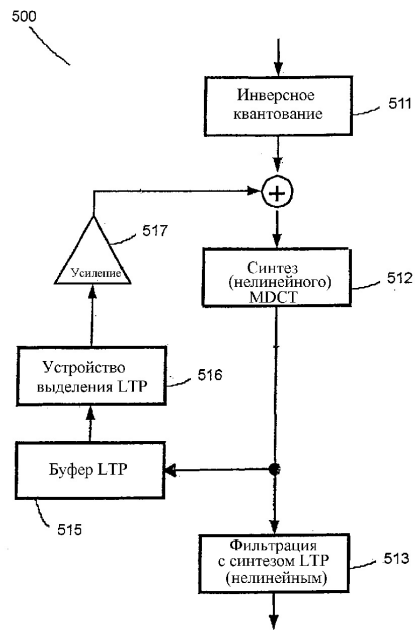
2



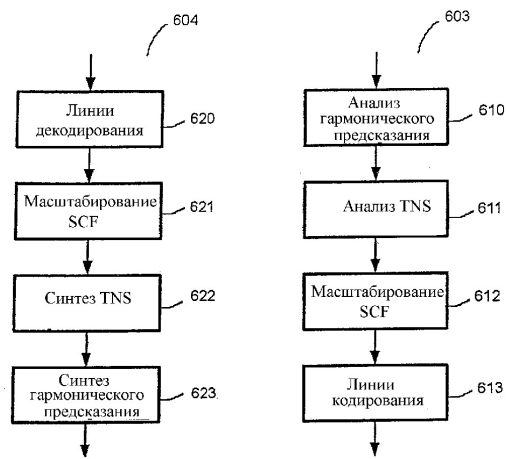
Фиг. 3



Фиг. 4



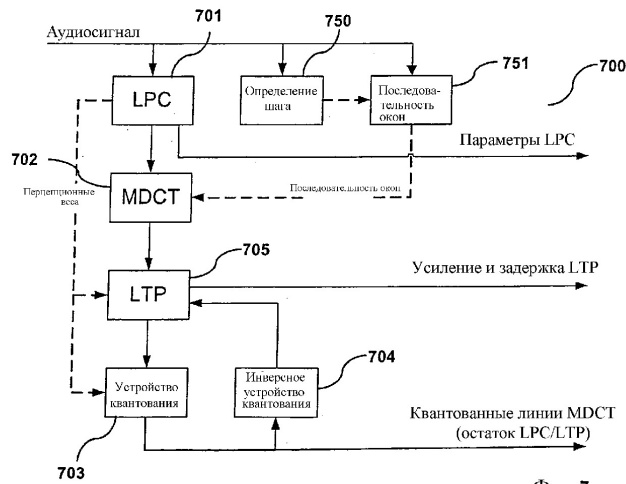
Фиг. 5



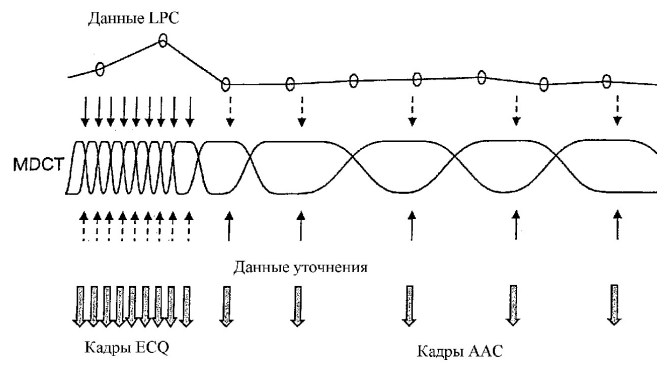
Фиг. 6



Фиг. 7

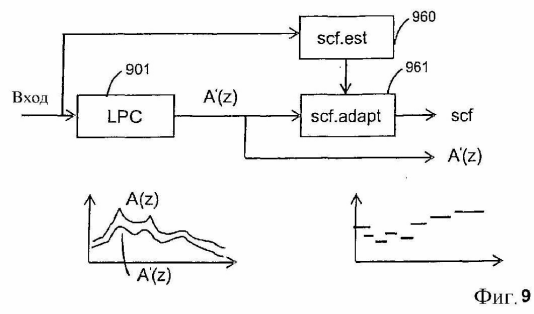


Фиг. 7а

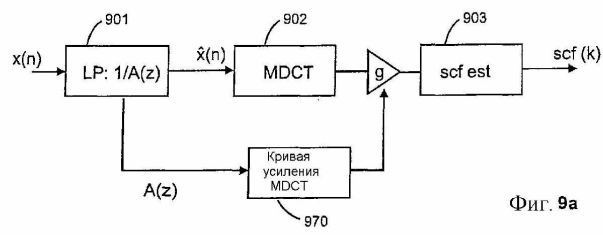


Фиг. 8

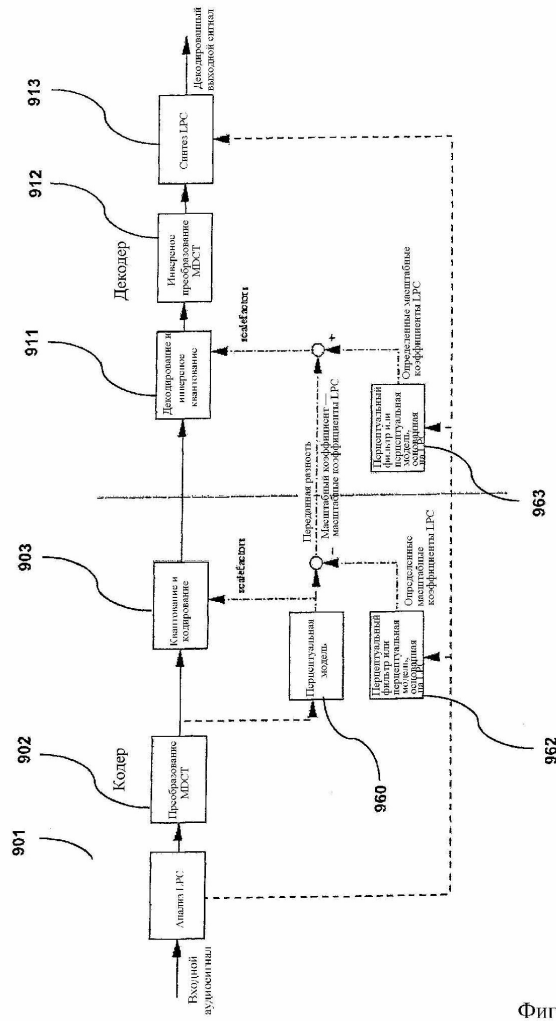
8/21



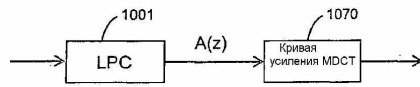
Фиг. 9



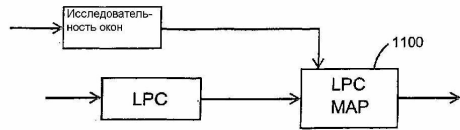
Фиг. 9а



Фиг. 9b

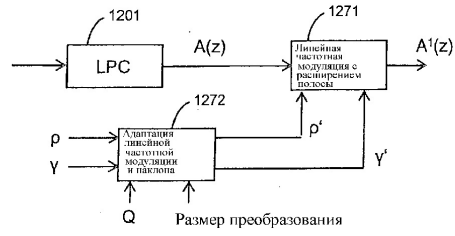


Фиг. 10



Фиг. 11

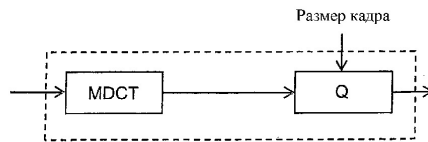
10/21



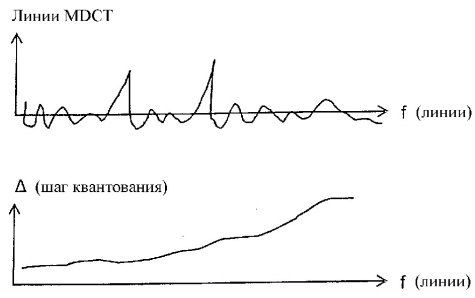
Фиг. 12



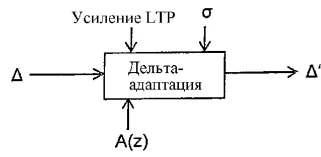
Фиг. 13



Фиг. 14



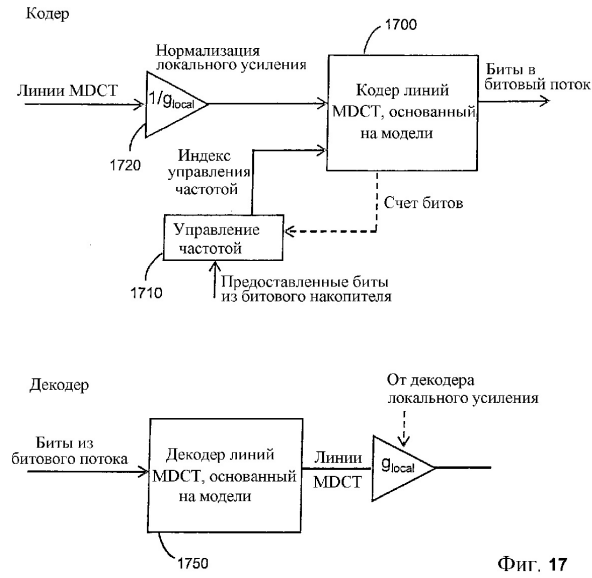
Фиг. 15



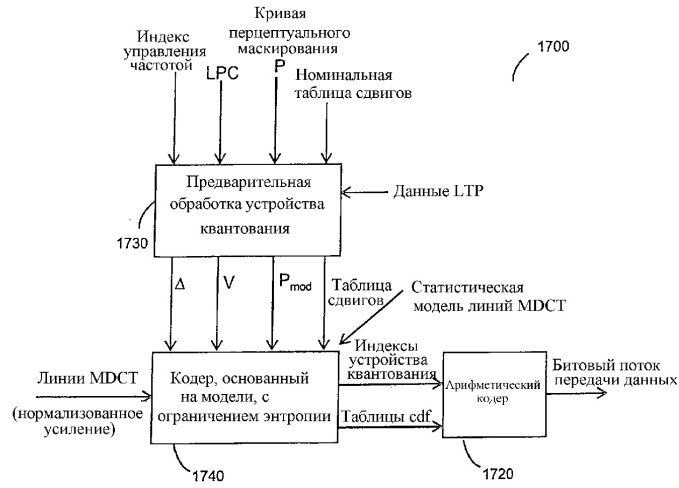
Фиг. 15а



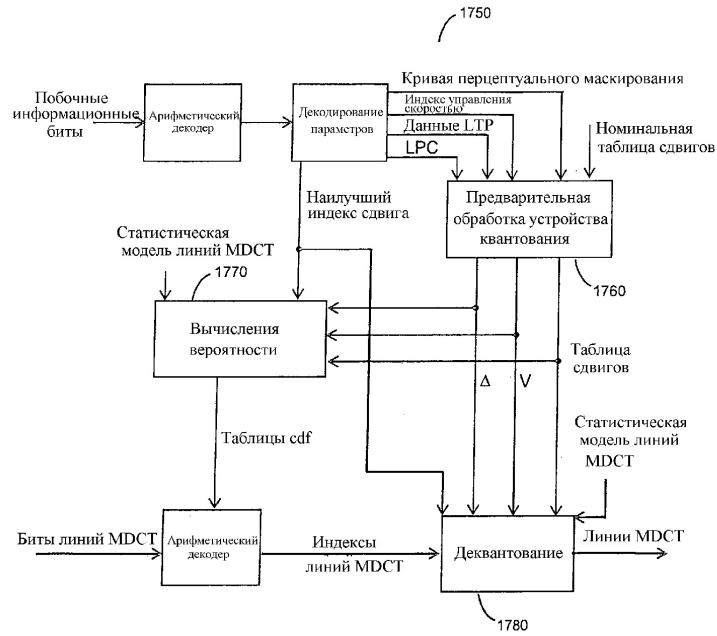
Фиг. 16



Фиг. 17

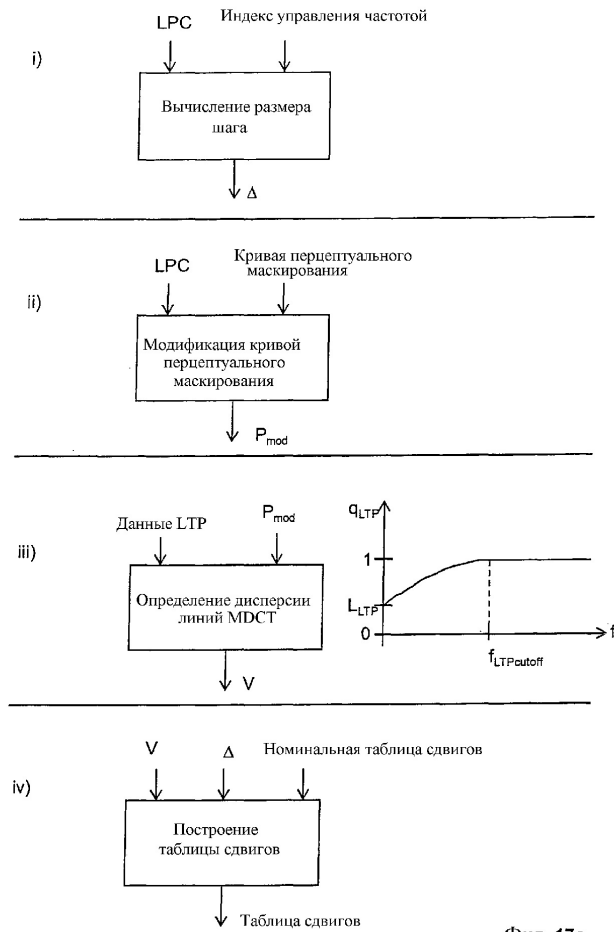


Фиг. 17а

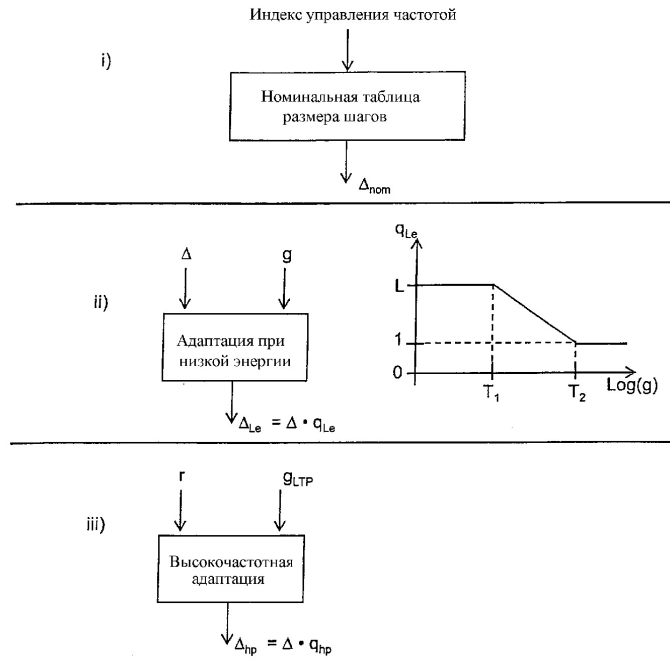


Фиг. 17b

15/21



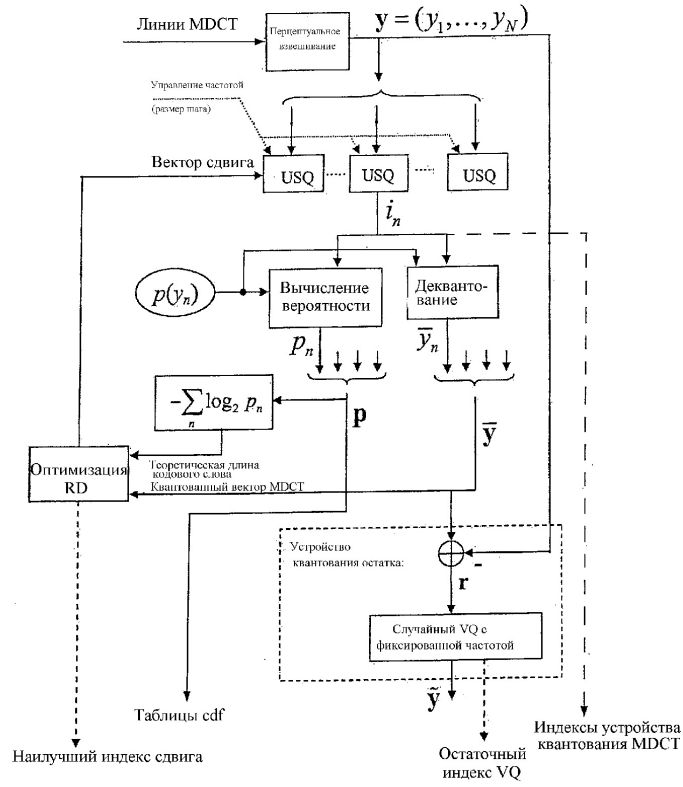
16/21



Фиг. 17d

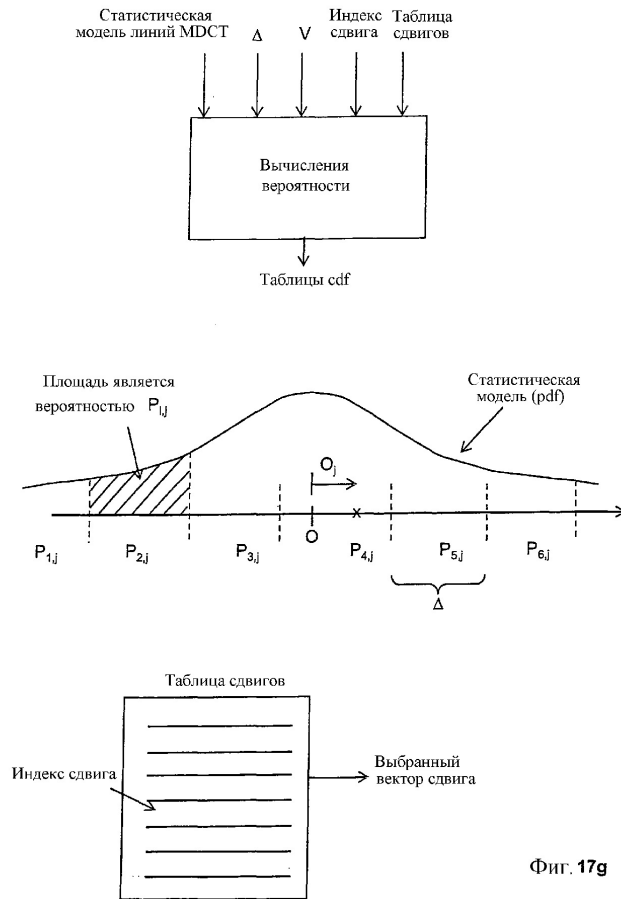


Фиг. 17f



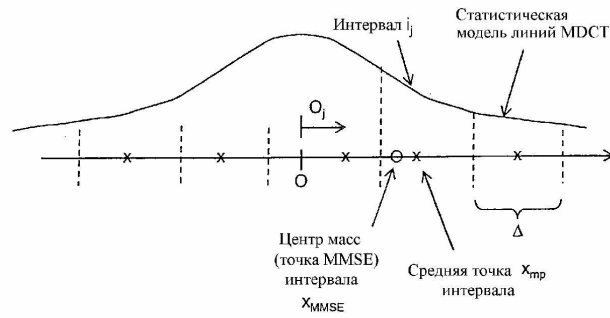
Фиг. 17е

18/21

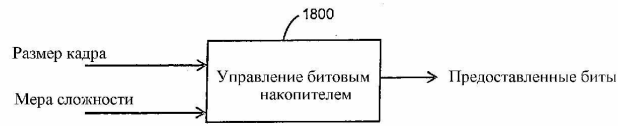


Фиг. 17g

19/21

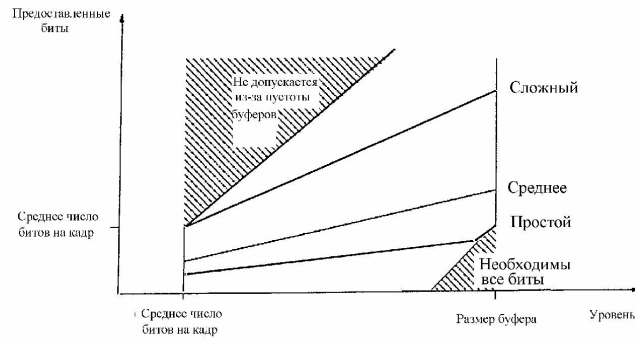


Фиг. 17h

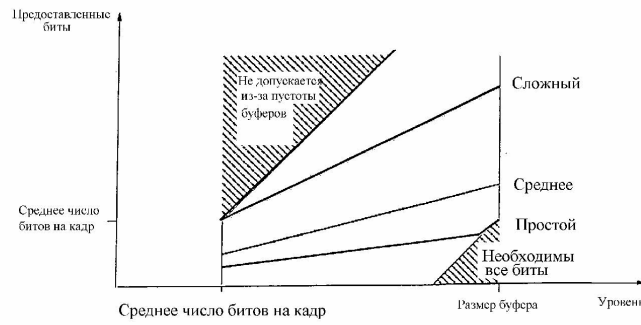


Фиг. 18

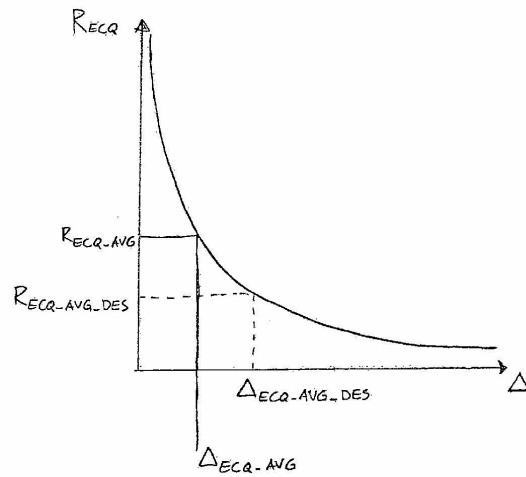
20/21



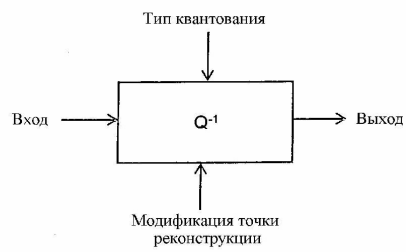
Фиг. 18а



Фиг. 18б



Фиг. 18с



Фиг. 19