



ФЕДЕРАЛЬНАЯ СЛУЖБА
ПО ИНТЕЛЛЕКТУАЛЬНОЙ СОБСТВЕННОСТИ

(12) ОПИСАНИЕ ИЗОБРЕТЕНИЯ К ПАТЕНТУ

(52) СПК

G10L 19/008 (2024.08); G10L 19/032 (2024.08)

(21)(22) Заявка: 2023100224, 10.06.2021

(24) Дата начала отсчета срока действия патента:
10.06.2021Дата регистрации:
15.04.2025

Приоритет(ы):

(30) Конвенционный приоритет:
11.06.2020 US 63/037,784;
27.05.2021 US 63/194,010

(45) Опубликовано: 15.04.2025 Бюл. № 11

(85) Дата начала рассмотрения заявки РСТ на
национальной фазе: 11.01.2023(86) Заявка РСТ:
US 2021/036886 (10.06.2021)(87) Публикация заявки РСТ:
WO 2021/252811 (16.12.2021)

Адрес для переписки:

123242, г. Москва, Кудринская площадь, д. 1,
а/я 35, Кузнецова Светлана Анатольевна

(72) Автор(ы):

МАКГРАТ, Дэвид С. (US),
ТЯГИ, Ришаб (US),
БРАУН, Стефани (US),
ТОРРЕС, Хуан Феликс (US)

(73) Патентообладатель(и):

ДОЛБИ ЛАБОРАТОРИС ЛАЙСЭНЗИН
КОРПОРЕЙШН (US)(56) Список документов, цитированных в отчете
о поиске: US 2012035941 A1, 09.02.2012.MCGRATH D., BRUHN S., PURNHAGEN H.,
ECKERT M., TORRES J., BROWN S., DARCY
D. Immersive audio coding for virtual reality
using a metadata-assisted extension of the 3GPP
EVS codec // 2019 IEEE International Conference
on Acoustics, Speech and Signal Processing
(ICASSP). doi:10.1109/icassp.2019.8683712, pp.730-
732, фиг. (см. прод.)(54) КВАНТОВАНИЕ И ЭНТРОПИЙНОЕ КОДИРОВАНИЕ ПАРАМЕТРОВ ДЛЯ АУДИОКОДЕКА
С НИЗКОЙ ЗАДЕРЖКОЙ

(57) Реферат:

Изобретение относится к области энтропийного кодирования параметров для аудиокодексов с низкой задержкой и к механизмам для достижения целевых битрейтов параметров путем итеративного уточнения битрейта параметров с использованием диапазона технологий квантования и энтропийного кодирования. Техническим результатом является минимизация дополнительной информации (параметров) при сохранении высокой частоты обновления кадров для основы аудиоданных. Заявлен способ покадрового кодирования метаданных для входного сигнала, причем

метаданные содержат множество по меньшей мере частично взаимосвязанных параметров, вычисляемых из входного сигнала. Способ включает для каждого кадра итеративное выполнение, с использованием зацикленного процесса, следующих этапов: определение стратегии обработки из множества стратегий обработки для вычисления и квантования параметров, вычисление и квантование параметров на основе определенной стратегии обработки для получения квантованных параметров, и кодирование квантованных параметров. В частности, каждая из множества

стратегий обработки содержит соответствующий первый указатель, указывающий порядок, относящийся к вычислению и квантованию отдельных параметров. Стратегию обработки

определяют на основе по меньшей мере одного порогового значения битрейта. 3 н. 14 з.п. ф-лы, 4 ил., 3 табл.



Фиг. 2

(56) (продолжение):

1. US 2008228476 A1, 18.09.2008. US 2019304475 A1, 03.10.2019. WO 2005008628 A1, 27.01.2005. RU 2621003 C2, 30.05.2017. RU 2606552 C2, 10.01.2017.

RU 2838373 C1

RU 2838373 C1



FEDERAL SERVICE
FOR INTELLECTUAL PROPERTY

(51) Int. Cl.
G10L 19/008 (2013.01)
G10L 19/032 (2013.01)

(12) **ABSTRACT OF INVENTION**

(52) CPC
G10L 19/008 (2024.08); *G10L 19/032* (2024.08)

(21)(22) Application: **2023100224, 10.06.2021**

(24) Effective date for property rights:
10.06.2021

Registration date:
15.04.2025

Priority:

(30) Convention priority:
11.06.2020 US 63/037,784;
27.05.2021 US 63/194,010

(45) Date of publication: **15.04.2025** Bull. № 11

(85) Commencement of national phase: **11.01.2023**

(86) PCT application:
US 2021/036886 (10.06.2021)

(87) PCT publication:
WO 2021/252811 (16.12.2021)

Mail address:

**123242, g. Moskva, Kudrinskaya ploshchad, d. 1,
a/ya 35, Kuznetsova Svetlana Anatolevna**

(72) Inventor(s):

**MCGRATH, David S. (US),
TYAGI, Rishabh (US),
BROWN, Stefanie (US),
TORRES, Juan Felix (US)**

(73) Proprietor(s):

**DOLBY LABORATORIES LICENSING
CORPORATION (US)**

(54) **QUANTISATION AND ENTROPY CODING OF PARAMETERS FOR LOW-LATENCY AUDIO CODEC**

(57) Abstract:

FIELD: data processing.

SUBSTANCE: invention relates to the field of entropy coding of parameters for audio codecs with low latency and to mechanisms for achieving target bitrates of parameters by iterative refinement of bitrate of parameters using a range of technologies of quantisation and entropy coding. Disclosed is a method for frame-by-frame encoding of metadata for an input signal, wherein the metadata contain a plurality of at least partially interrelated parameters calculated from the input signal. Method involves iterative execution of the following steps for each frame using a looped process: determining a processing strategy from a plurality of processing strategies for calculating and quantizing

parameters, calculating and quantizing parameters based on the determined processing strategy to obtain quantized parameters, and encoding the quantized parameters. In particular, each of the plurality of processing strategies comprises a corresponding first indicator indicating the order related to the calculation and quantization of the individual parameters. Processing strategy is determined based on at least one bitrate threshold value.

EFFECT: minimizing additional information (parameters) while maintaining a high frame refresh rate for the audio data base.

17 cl, 4 dwg, 3 tbl

RU 2 838 373 C1

RU 2 838 373 C1



Фиг. 2

ПЕРЕКРЕСТНЫЕ ССЫЛКИ НА РОДСТВЕННЫЕ ЗАЯВКИ

Данная заявка заявляет приоритет предварительных заявок на патент США № 63/037,784 и 63/194,010, поданных 11 июня 2020 г. и 27 мая 2021 г. соответственно, каждая из которых включена посредством ссылки в своей полноте.

ОБЛАСТЬ ТЕХНИКИ, К КОТОРОЙ ОТНОСИТСЯ ИЗОБРЕТЕНИЕ

Настоящее изобретение в общем относится к области энтропийного кодирования параметров (дополнительная информация) для аудиокодеков с низкой задержкой (кодеров/декодеров) и к механизмам для достижения целевых битрейтов параметров путем итеративного уточнения битрейта параметров с использованием диапазона технологий квантования и энтропийного кодирования.

УРОВЕНЬ ТЕХНИКИ

Когда период кадра (размер кадра) аудиокодека (кодера/декодера) приближается к 20 миллисекундам (мс) или меньше, основу аудиоданных обновляют в коротких размерах кадра. Если следовать подходу обновления как основы аудиоданных, так и параметров каждый кадр, дополнительная информация для каждого кадра также будет встраиваться и передаваться с той же частотой.

Однако в данной области техники в общем известно, что дополнительную информацию не нужно обновлять настолько часто. Например, пространственные параметры обычно могут вычисляться и обновляться, например, каждые 40 мс. Для кодеков с периодами кадра, равными 40 мс или более, это обычно означает, что частота обновления параметров соответствует частоте кадров, и таким образом параметры могут быть закодированы в каждом кадре независимо. Однако в кодеках с короткими периодами кадра, например менее 40 мс, это означает, что параметры будут фактически чрезмерно дискретизированы, если все они будут включаться в каждом кадре.

Таким образом, в широком смысле, цель настоящего изобретения заключается в предоставлении механизмов для того, чтобы насколько возможно свести к минимуму дополнительную информацию (которую иногда также называют параметрами), но сохранить высокую частоту обновления кадров для основы аудиоданных.

СУЩНОСТЬ ИЗОБРЕТЕНИЯ

Учитывая вышеуказанное, настоящее изобретение в общем предоставляет способ покадрового кодирования метаданных для входного сигнала, а также соответствующие программу, машиночитаемый носитель данных и устройство, обладающие признаками, указанными в соответствующих независимых пунктах формулы изобретения.

Согласно одному аспекту настоящего изобретения предоставлен способ покадрового кодирования метаданных для входного сигнала. В частности, метаданные можно рассчитать или вычислить (например, извлечь) из входного (звукового или видео) сигнала с использованием подходящего кодека (кодера/декодера). В общем говоря, метаданные можно использовать для восстановления входного сигнала на стороне декодера. Метаданные могут содержать множество по меньшей мере частично взаимосвязанных параметров, вычисляемых из входного сигнала. Другими словами, по меньшей мере некоторые параметры входного сигнала могут быть вычислены (например, сгенерированы или восстановлены) в зависимости от по меньшей мере некоторых других параметров, так что, в зависимости от различных обстоятельств, не все параметры необходимо всегда передавать явным образом.

В частности, способ может содержать/включать, для каждого кадра, итеративное выполнение, с использованием зацикленного процесса, следующих этапов: определение стратегии обработки из множества стратегий обработки для вычисления и квантования параметров; вычисление и квантование параметров на основе определенной стратегии

обработки для получения квантованных параметров; и кодирование квантованных параметров. Поскольку зацикленный процесс в общем направлен (помимо прочего) на обработку, связанную с квантованием, в некоторых случаях зацикленный процесс также может называться циклом квантования (или просто циклом, для краткости).

5 Подобным образом, поскольку стратегия обработки также в общем направлена (помимо прочего) на обработку, которая связана с квантованием, в некоторых случаях стратегия обработки также может называться стратегией квантования (или, в некоторых других случаях, взаимозаменяемо называться схемой квантования). Кроме этого необходимо отметить, что процесс кодирования может использовать любую подходящую процедуру

10 кодирования, включая, но без ограничения, энтропийное кодирование (например, кодирование Хаффмана или арифметическое кодирование) или кодирование без энтропии (например, кодирование с основанием 2). В зависимости от различных реализаций и/или требований может использоваться любой другой подходящий механизм кодирования.

15 Как будет понятно и очевидно специалисту в данной области, множество стратегий обработки для вычисления и квантования параметров могут быть предоставлены любым подходящим образом, например предварительно определенным или предварительно сконфигурированным. Соответственно, стратегия обработки также может быть определена из множества стратегий обработки любым подходящим образом.

20 Например, в зависимости от (текущего) требования к битрейту, подходящая стратегия обработки может быть выбрана из множества стратегий обработки, так что итоговый битрейт после выполнения вычисления, квантования и кодирования (например, с энтропийным кодированием или без него) на основе выбранной таким образом стратегии обработки удовлетворяет (текущему) требованию к битрейту. Следует отметить, что

25 поскольку требование к битрейту может изменяться время от времени (например, от кадра к кадру), стратегия обработки, определенная таким образом, также может быть разной для каждого кадра или некоторых кадров.

В частности, каждая из множества стратегий обработки может содержать соответствующий первый указатель, указывающий порядок (или последовательность),

30 относящуюся к вычислению и квантованию отдельных параметров. Другими словами, первый указатель может содержать информацию о последовательности, указывающую, когда и в каком порядке выполнять вычисление и квантование отдельных параметров. В качестве примера (но не ограничения), первый указатель может содержать информацию, указывающую, что все параметры сначала вычисляют, перед тем как

35 выполнять квантование любого из них.

Более конкретно, стратегию обработки определяют на основе по меньшей мере одного порогового значения битрейта. Как будет понятно и очевидно специалисту в данной области, пороговые значение(-я) битрейта могут быть, например, предварительно определены или предварительно сконфигурированы, в зависимости от различных

40 реализаций и/или требований.

Говоря в общем, в вышеописанной конфигурации предложенный способ согласно настоящему изобретению можно рассматривать как внедряющий идею итеративного и поэтапного подхода к выбору оптимальной схемы/стратегии квантования параметров, который в общем ищет «наилучшую» (или оптимальную) схему квантования из

45 нескольких альтернатив. Тем не менее, следует отметить, что в настоящем случае термин «наилучшая» не обязательно должен относиться к схеме квантования с наименьшим (итоговым) битрейтом параметров (т.е. после квантования и возможного кодирования), но может относиться к схеме, способной облегчить потерю состояния для декодера. Как

будет понятно специалисту, в общем «состояние» декодера относится к истории информации, которую декодер сохраняет из предыдущих кадров для того, чтобы иметь возможность правильно декодировать текущий кадр. Например (но не в качестве ограничения), в некоторых случаях сторона кодера может применять так называемое временно-дифференциальное кодирование. Однако использование временно-дифференциального кодирования в общем может обладать недостатком, который в первую очередь заключается в том, что обычно вводится состояние от кадра к кадру, что может стать причиной проблем, когда во время передачи звуковой поток может столкнуться с потерей пакетов. В этом случае как звуковые данные, так и параметры, связанные со звуковыми данными, могут быть утеряны во время передачи, так что любые параметры, которые были обновлены с помощью временно-дифференциального кодирования, могут столкнуться с потенциальными артефактами в нескольких последующих кадрах. В этом смысле вышеупомянутое облегчение потери состояния относится к попытке избежать временно-дифференциального кодирования, где это возможно, так что декодеру не нужно полагаться на метаданные, принятые в предыдущих кадрах, чтобы декодировать метаданные текущего кадра. Однако когда требуется временно-дифференциальное кодирование, оно должно быть выполнено таким образом, чтобы система быстро восстановилась от потери пакетов. В частности, путем тщательного выбора подходящей схемы квантования, как описано в настоящем изобретении, представленное выше нежелательное поведение, связанное с потерей пакетов, можно насколько возможно ограничить (уменьшить). Другими словами, настоящее изобретение в общем предлагает уменьшение кодирования (стороны кодера), что включает итеративный процесс выбора для квантования и кодирования (с энтропией или без нее), который пытается свести к минимуму степень возможного внесения артефактов, вызванных потерей пакетов, например из-за использования временно-дифференциального кодирования.

В некоторых примерах стратегия обработки может быть определена так, что (итоговый) битрейт закодированных квантованных параметров равен пороговому значению битрейта (метаданных/параметров) или меньше него. По существу, итоговый битрейт после квантования и кодирования с использованием определенной (например, выбранной) стратегии обработки находится в рамках (по меньшей мере одного) порогового значения битрейта, тем самым удовлетворяя требованию к битрейту, например утвержденному заранее или предварительно определенному техническими условиями стандартизации.

В некоторых примерах каждая из множества стратегий обработки может дополнительно содержать соответствующий второй указатель, указывающий информацию для выполнения квантования параметров.

В некоторых примерах информация для выполнения квантования параметров содержит соответствующие диапазоны квантования и/или уровни квантования для множества параметров. Например, информация может относиться к максимальному значению, минимальному значению, количеству уровней квантования или любому другому подходящему значению, требуемому для каждого из соответствующих параметров (например, к одному соответствующему параметру на каждый тип параметров). В общем говоря, как будет понятно и очевидно специалисту в данной области, эти связанные с квантованием значения/параметры предоставляют или определяют более грубое или более точное квантование в целом и соответственно сопутствующее лучшее или худшее пространственное воспроизведение. Как будет понятно и очевидно специалисту в данной области, вообще говоря, некоторые

параметры (квантования) в общем считаются более чувствительными к квантованию, чем другие, и в общем может отсутствовать абсолютная методика точного/грубого квантования для всех параметров.

В вышеописанной конфигурации можно считать, что каждая из множества стратегий обработки содержит первый указатель (его первую часть/порцию), относящийся к порядку/последовательности, связанным с вычислением и квантованием; и второй указатель (его вторую часть/порцию), относящийся к фактическому процессу квантования. Путем тщательного проектирования стратегии обработки (например, с помощью разных комбинаций первого указателя и второго указателя) можно эффективно и гибко добиться соответствия разным конфигурациям/требованиям битрейта, например для разных случаев или сценариев использования. В частности, в некоторых случаях может существовать одна стратегия обработки (например, стратегия наиболее грубого квантования из множества стратегий квантования), которая может считаться гарантирующей битрейт ниже целевого порогового значения битрейта (или равный ему).

В некоторых примерах кодирование параметров может включать временно- и/или частотно-дифференциальное кодирование. Говоря в общем, один параметр метаданных можно квантовать из непрерывного числового значения в индекс, представляющий собой дискретное значение. При недифференциальном кодировании информация, кодируемая для этого параметра метаданных, соответствует непосредственно этому индексу. Следует отметить, что термин «недифференциальное кодирование», используемый в настоящем изобретении, может относиться к не являющемуся временно-дифференциальным кодированию, не являющемуся частотно-дифференциальным кодированию или не являющемуся дифференциальным кодированием всех видов, если это необходимо, как будет понятно и очевидно специалисту в данной области. При временно-дифференциальном кодировании кодируемая информация представляет собой разницу между индексом параметра метаданных из текущего кадра и индексом того же параметра метаданных из предыдущего кадра. Как будет понятно и очевидно специалисту в данной области, вышеописанная общая идея временно-дифференциального кодирования может быть дополнительно расширена, например на множество полос частот. Соответственно, параметр метаданных может быть расширен подобным образом, например на множество параметров, соответственно относящихся ко множеству полос частот (или каждой из них), как необходимо. Частотно-дифференциальное кодирование следует подобному принципу, но закодированная разница является разницей между метаданными одной полосы частот текущего кадра и метаданными другой полосы частот текущего кадра (в отличие от текущего кадра минус предыдущий кадр во временно-дифференциальном кодировании). В качестве простого примера (но не ограничения), предположим, что a_0 , a_1 , a_2 и a_3 будут обозначать индексы параметров в 4 полосах частот конкретного кадра, тогда в одном примере реализации частотно-дифференциальные индексы могут представлять собой a_0 , a_0-a_1 , a_1-a_2 , a_2-a_3 . Как будет очевидно специалисту в данной области, общая идея (временно-и/или частотно-) дифференциального кодирования заключается в том, что метаданные обычно могут медленно изменяться от кадра к кадру или от полосы частот к полосе частот, так что даже если исходное значение метаданных было большим, разница между ним и метаданными предыдущего кадра или разница между ним и метаданными другой полосы частот вероятно будет небольшой. Это является преимуществом, поскольку в общем параметры со статистическими распределениями, которые стремятся к нулю, могут быть закодированы с использованием меньшего количества битов.

В некоторых примерах стратегия обработки, определенная для текущего кадра, может отличаться от стратегии обработки, определенной для предыдущего кадра, и соответственно кодирование параметров может включать временно-дифференциальное кодирование в разных стратегиях обработки. Другими словами, в определенных случаях, где определены разные стратегии обработки (например, для разных кадров входного сигнала), способ согласно настоящему изобретению по-прежнему может кодировать параметры, например с использованием временно-дифференциального кодирования в этих разных стратегиях обработки.

Как указано выше, каждая из множества стратегий обработки может содержать соответствующий первый указатель, указывающий порядок (или последовательность), относящийся к вычислению и квантованию отдельных параметров.

В некоторых примерах первый указатель может содержать информацию, указывающую, что все параметры вычисляются перед квантованием.

В некоторых примерах первый указатель может содержать информацию, указывающую, что параметры вычисляются по отдельности и затем квантуют последовательно друг за другом. В частности, по меньшей мере один параметр из множества параметров может быть вычислен на основе другого квантованного параметра из множества параметров. В качестве примера, но не ограничения, предположим, что всего три параметра необходимо вычислить и квантовать, тогда первый параметр можно вычислить в первую очередь (из входного сигнала) и затем квантовать; тогда как второй параметр можно вычислить на основе (квантованного) первого параметра и затем квантовать сам второй параметр; и наконец, третий параметр вычисляют на основе (квантованного) первого параметра и/или (квантованного) второго параметра и затем квантуют. В одном примере третий параметр вычисляют на основе квантованных первого и второго параметров.

В некоторых примерах первый указатель может содержать информацию, указывающую, что все параметры вычисляются перед квантованием любого параметра; и в частности, по меньшей мере один из параметров повторно вычисляют на основе другого квантованного параметра и повторно вычисленный параметр квантуют. Также принимая в качестве примера вышеописанные три параметра, все параметры сначала вычисляют и затем квантуют первый и второй параметры; после этого повторно вычисляют третий параметр, например на основе квантованных вторых параметров, и затем третий параметр квантуют на основе повторно вычисленного значения.

В некоторых примерах способ может дополнительно включать, перед кодированием квантованных параметров, отображение индексов квантованных параметров из предыдущего кадра в индексы текущего кадра. Другими словами, если определяют (например, находят/выбирают) другую стратегию обработки (схему квантования, например, что касается разных уровней и/или последовательностей квантования), индексы (квантования) из предыдущего кадра, которые были квантованы с использованием другой схемы квантования, отображают в индексы текущего кадра. Следует отметить, что это позволяет выполнять временно-дифференциальное кодирование между кадрами без необходимости отправлять недифференциальный кадр при каждом изменении схемы квантования, тем самым дополнительно улучшая общую эффективность и гибкость кодирования.

В некоторых возможных реализациях отображение индексов может быть выполнено на основе формул:

$$index_{cur} = round(index_{prev} \times (quant_lvl_{cur} - 1) / (quant_lvl_{prev} - 1)), \text{ где } index_{cur}$$

обозначает индексы текущего кадра после отображения, $index_{prev}$ обозначает индекс
 5 предыдущего кадра, $quant_lvl_{cur}$ обозначает уровень квантования текущего кадра и
 $quant_lvl_{prev}$ обозначает уровень квантования предыдущего кадра.

В качестве простого иллюстративного примера предположим, что диапазон
 10 квантования составляет от 0 до 2, и предположим, что предыдущие уровни квантования
 составляют 11. В случае равномерного квантования это в общем означало бы, что
 каждый этап квантования составлял бы 0,2. Далее предположим, что текущие уровни
 квантования составляют 21, что означает, что каждый этап квантования составляет
 15 0,1 при равномерном квантовании. На основе этих предположений, если квантованное
 значение в предыдущем кадре составляло 0,4, то при 11 равномерных уровнях
 квантования получают следующий предыдущий индекс $index_{prev} = 2$. Отображение
 предоставляет квантованные индексы метаданных предыдущего кадра, как если бы
 они были квантованы с использованием уровней квантования текущего кадра. Таким
 20 образом, в этом примере, если уровни квантования в текущем кадре составляют 21, то
 квантованное значение 0,4 будет отображено в $index_{curr} = 4$. После того, как

отображенные индексы вычислены, вычисляют разницу между индексами текущего
 25 кадра и предыдущего кадра и кодируют эту разницу. Аналогичные или подобные
 подходы при необходимости также могут применяться к частотно-дифференциальному
 кодированию, как будет понятно и очевидно специалисту в данной области.

Необходимо отметить, что вышеописанные формулы и соответствующий пример
 30 предоставлены лишь с иллюстративной целью и что любой другой подходящий механизм
 (например, справочная таблица и т.д.) может использоваться для выполнения
 отображения индексов, как будет понятно и очевидно специалисту в данной области.

В некоторых примерах по меньшей мере одно пороговое значение битрейта может
 содержать целевое пороговое значение битрейта. Соответственно, зацикленный процесс
 может включать следующие этапы: квантование и кодирование параметров
 35 недифференциальным и/или частотно-дифференциальным образом с использованием
 энтропийного кодера в соответствии с (определенной) стратегией обработки; оценивание
 (например, вычисление) первого битрейта параметров для закодированных параметров;
 и выход из зацикленного процесса, если первый битрейт параметров меньше целевого
 порогового значения битрейта или равен ему. В частности, в некоторых возможных
 40 реализациях первый битрейт параметров можно оценить (вычислить) из минимума
 среди схем недифференциального и частотно-дифференциального кодирования,
 закодированных с использованием (обученных) энтропийных кодеров. Как будет
 понятно и очевидно специалисту в данной области, энтропийные кодеры могут быть
 обучены любым подходящим образом, например для приспособления к отдельным
 45 схемам кодирования. Например, в некоторых возможных реализациях обучение
 энтропийных кодеров может включать разработку вероятностных моделей на основе
 метаданных, вычисленных из большой группы входных сигналов. Ожидается, что
 конкретные сигналы, выбранные для разработки этих моделей, будут представлять

типы сигналов, которые, как ожидается, будут проходить через систему при повседневном использовании. Таким образом, метаданные из других подобных сигналов должны быть закодированы насколько можно более эффективно. Вкратце, если говорить в общем, целью этого обучения является адаптация энтропийных кодеров для придания им максимальной эффективности с ожидаемым распределением вероятностей параметров.

В некоторых примерах, зацикленный процесс может дополнительно включать следующие этапы: если первый битрейт параметров больше целевого порогового значения битрейта, квантование и кодирование параметров недифференциальным образом без энтропии согласно стратегии обработки; оценивание второго битрейта параметров для закодированных параметров; и выход из зацикленного процесса, если второй битрейт параметров меньше целевого порогового значения битрейта или равен ему.

В некоторых примерах зацикленный процесс может дополнительно включать следующие этапы: если второй битрейт параметров больше целевого порогового значения битрейта, квантование и кодирование параметров временно-дифференциальным образом с использованием (обученного) энтропийного кодера согласно стратегии обработки; оценивание третьего битрейта параметров для закодированных параметров; и выход из зацикленного процесса, если третий битрейт параметров меньше целевого порогового значения битрейта или равен ему.

В некоторых примерах временно-дифференциальное квантование и кодирование может выполняться на подмножестве параметров методом с чередованием частот относительно предыдущего кадра. В частности, как будет понятно и очевидно специалисту в данной области, метод с чередованием частот обычно может относиться к случаям, где разные полосы частот (например, соответствующие разным подмножествам параметров) обрабатываются (например, подвергаются квантованию и кодированию) для разных кадров. Другими словами, временно-дифференциальное квантование и кодирование (по меньшей мере подмножества) параметров для текущего кадра могут быть выполнены в другой полосе частот (соответствующей обрабатываемым в настоящее время параметрам), которая отличается от полосы частот предыдущего кадра.

В некоторых примерах временно-дифференциальное квантование и кодирование могут быть выполнены путем циклического прохождения через ряд схем временно-дифференциального кодирования с чередованием частот таким образом, что для каждого цикла другое подмножество параметров (соответствующее другому множеству полос частот) квантуют и кодируют временно-дифференциально, тогда как остальные параметры квантуют и кодируют недифференциально.

В некоторых примерах определенная стратегия обработки может считаться первой стратегией обработки, и соответственно зацикленный процесс может дополнительно включать следующие этапы: если третий битрейт параметров больше целевого порогового значения битрейта, определение второй стратегии обработки из множества стратегий обработки, так, что ожидается, что (итоговый) битрейт после применения второй стратегии обработки будет меньше битрейта использования первой стратегии обработки; и повторение вышеописанных этапов зацикленного процесса. Как будет понятно и очевидно специалисту в данной области, в таких случаях определенная (например, выбранная) таким образом вторая стратегия обработки может просто считаться стратегией обработки, которая является более грубой, чем определенная (например, выбранная) ранее первая стратегия обработки. По существу, размер

множества возможных квантованных значений/индексов можно уменьшить, что (обычно) также приводит к соответственному уменьшению битрейта.

В некоторых примерах параметры могут быть представлены в первом количестве полос частот, и зацикленный процесс может дополнительно включать следующие этапы: если третий битрейт параметров больше целевого порогового значения битрейта, уменьшение количества полос частот, представляющих параметры, до второго количества, которое меньше первого количества, так что общее количество параметров, которые необходимо квантовать и кодировать, уменьшается; и повторение вышеописанных этапов зацикленного процесса.

В некоторых примерах параметры представлены в первом количестве полос частот, и зацикленный процесс может дополнительно включать следующие этапы: если третий битрейт параметров больше целевого порогового значения битрейта, то применяют повторное использование (которое в некоторых случаях называют «замораживанием») параметров в одной или более полос частот из предыдущего кадра в текущем кадре; и повторение этапов вышеописанного зацикленного процесса. В качестве примера, при кодировании с использованием конкретной схемы кодирования можно замораживать параметры в определенных полосе(-ах) частот (например, частотных полосах 2, 6 и 10). В качестве дополнительного иллюстративного примера, если все полосы частот замораживают в течение периода 2 кадров, то кодер может передавать половину полос (например, полосы с четными номерами) в кадре N, а остальную половину (например, полосы с нечетными номерами) в кадре N+1 (тем самым уменьшая общее количество параметров, которые необходимо передавать), что в общем означает, что декодер будет получать все (например, 12) обновленных полос частот раз в два кадра. В таких случаях, если один кадр утерян, обычно остается возможность экстраполяции из двух последних пригодных кадров. При восстановлении после потери пакетов можно выполнить интерполирование между полосами, которые были приняты с заданным кадром. В общем говоря, результатом вышеописанного процесса замораживания будет уменьшение энтропии, что не требует изменения декодера или схемы энтропийного кодирования, с небольшим влиянием на качество.

В итоге, когда необходимо уменьшить общее количество полос, это можно выполнить по меньшей мере двумя следующими методами. Первый метод заключается в уменьшении разрешающей способности по частоте, при котором вместо использования N полос используются только M полос (где $M < N$), и ширина одной или более полос в конфигурации с M полосами выше, чем в конфигурации с N полосами. Эти M полос могут быть производными от N полос, например смежные полосы могут быть сгруппированы вместе попарно, по три и так далее, или в виде любых других групп, которые имеют значимость для восприятия. Второй метод заключается в уменьшении разрешающей способности по времени, при котором значения ширины всех N полос могут оставаться в точности неизменными в частотной области, но полосы замораживаются в течение периода x кадров (где $x > 1$). Это означает, что обновления для N полос могут быть переданы в течение периода x кадров или, другими словами, только N/x полос из N полос необходимо обновлять и передавать на декодер с каждым кадром.

В некоторых примерах по меньшей мере одно пороговое значение битрейта может дополнительно содержать, в дополнение к вышеописанному целевому пороговому значению битрейта, максимальное пороговое значение битрейта, которое больше целевого порогового значения битрейта. Соответственно, зацикленный процесс может дополнительно включать следующие этапы: перед определением второй стратегии

обработки, или уменьшением количества полос частот, или повторным использованием параметров, получение минимума среди первого, второго и третьего битрейтов параметров; и выход из заикленного процесса, если минимум меньше максимального порогового значения битрейта или равен ему.

5 Стоит отметить, что если цикл обработки завершается конкретным этапом, как описано выше, то в общем это означает, что окончательный битрейт параметров является битрейтом, который вычислен на этом этапе (т.е. при выходе из цикла обработки). Кроме того, как указано выше, для большей верности может существовать определенная (например, самая грубая) стратегия квантования среди заданных стратегий квантования, доступных для квантования параметров, которая гарантированно 10 обеспечивает битрейт, который меньше целевого порогового значения битрейта или максимального порогового значения битрейта или равен им. Таким образом, можно гарантировать постоянное наличие решения для подходящего битрейта параметров в рамках целевого порогового значения битрейта или максимального порогового значения битрейта. 15

В некоторых примерах параметры могут содержать один или более параметров предсказания (иногда просто называемых параметрами PR), параметров перекрестного предсказания (иногда просто называемых параметрами C) и параметров декорреляции (иногда просто называемых параметрами P). Как указано выше, по меньшей мере 20 некоторые из параметров по меньшей мере частично взаимосвязаны, так что они могут быть вычислены на основе друг друга. Разумеется, как будет понятно и очевидно специалисту в данной области, могут существовать любые другие параметры (типы параметров), в зависимости от различных реализаций и/или требований (например, использования конкретных кодеков).

25 Как указано выше, порядок (или последовательность) вычисления и квантования параметров могут быть указаны первым указателем стратегий обработки.

В некоторых примерах параметры предсказания можно вычислять и квантовать в первую очередь, параметры перекрестного предсказания вычисляют из квантованных параметров предсказания и затем квантуют, а параметры декорреляции сначала 30 вычисляют из квантованных параметров перекрестного предсказания и квантованных параметров предсказания и затем квантуют.

В некоторых примерах параметры (т.е. параметры предсказания, параметры перекрестного предсказания и параметры декорреляции) могут быть сначала вычислены, затем параметры декорреляции и параметры предсказания квантуют, а параметры 35 перекрестного предсказания повторно вычисляют из квантованных параметров предсказания и затем квантуют.

В некоторых примерах способ может быть применен к кодированию метаданных кодека голосовых и звуковых служб с эффектом присутствия (IVAS) или кодека амбиофонии. Кодек амбиофонии может представлять собой кодек амбиофонии первого 40 порядка (FOA) или даже кодек амбиофонии более высокого порядка (HOA). Разумеется, как будет понятно и очевидно специалисту в данной области, могут применяться любые другие подходящие кодеки, в зависимости от различных реализаций.

В некоторых примерах размер кадра меньше 40 мс, и в частности равен или меньше 20 мс.

45 Согласно другому аспекту настоящего изобретения предоставлено устройство, содержащее процессор и память, соединенную с процессором. Процессор может быть приспособлен для вызова выполнения устройством всех этапов примерных способов, описанных в настоящем изобретении.

Согласно дополнительному аспекту настоящего изобретения предоставлена компьютерная программа. Компьютерная программа может включать команды, которые при исполнении процессором приводят к выполнению процессором всех этапов иллюстративных способов, описанных в настоящем изобретении.

5 Согласно еще одному аспекту предоставлен машиночитаемый носитель данных. Машиночитаемый носитель данных может хранить вышеупомянутую компьютерную программу.

Будет понятно, что признаки устройства и этапы способа можно взаимно заменять многими методами. В частности, детали раскрытых способа(-ов) могут быть
10 реализованы соответствующим устройством (или системой) и наоборот, как будет понятно специалисту. Кроме того, понятно, что любое из приведенных выше утверждений в отношении способа(-ов) также применимо и к соответствующему устройству (или системе) и наоборот.

КРАТКОЕ ОПИСАНИЕ ГРАФИЧЕСКИХ МАТЕРИАЛОВ

15 Примерные варианты осуществления настоящего изобретения описаны ниже со ссылкой на сопроводительные графические материалы, на которых

фиг. 1 представляет собой схематическое изображение структурной схемы кодера/декодера («кодека») для кодирования и декодирования сигналов (битовых потоков) согласно одному варианту осуществления настоящего изобретения,

20 фиг. 2 представляет собой блок-схему, изображающую пример способа покадрового кодирования метаданных для входного сигнала согласно одному варианту осуществления настоящего изобретения,

фиг. 3 представляет собой блок-схему, изображающую пример цикла обработки согласно одному варианту осуществления настоящего изобретения, и

25 фиг. 4 представляет собой блок-схему, изображающую пример цикла обработки согласно другому варианту осуществления настоящего изобретения.

ПОДРОБНОЕ ОПИСАНИЕ ИЗОБРЕТЕНИЯ

30 Фигуры (фиг.) и следующее описание относятся к предпочтительным вариантам осуществления только в качестве иллюстрации. Следует отметить, что из следующего обсуждения альтернативные варианты осуществления конструкций и способов, описанных в настоящем документе, будут легко понятными в качестве подходящих альтернатив, которые могут использоваться без отступления от заявленных принципов.

Теперь будет сделана подробная ссылка на несколько вариантов осуществления, примеры которых проиллюстрированы на прилагаемых фигурах. Нужно отметить,
35 что там, где это применимо, подобные или схожие ссылочные номера могут использоваться в фигурах и могут обозначать подобную или схожую функциональность. Фигуры изображают варианты осуществления описанной системы (или способа) только в целях иллюстрации. Специалисту в данной области будет легко понятно из следующего описания, что альтернативные варианты осуществления структур и способов,
40 изображенных в настоящем документе, могут применяться без отступления от принципов, заявленных в настоящем документе.

Кроме того, на фигурах, где соединительные элементы, такие как сплошные или пунктирные линии или стрелки, используют для иллюстрации соединения, отношения или связи между двумя или более другими схематическими элементами или среди них,
45 отсутствие любого из таких соединяющих элементов не подразумевает, что соединение, отношение или связь не может существовать. Другими словами, некоторые соединения, отношения или связи между элементами не показаны на графических материалах, чтобы не затруднять понимание настоящего изобретения. Дополнительно для упрощения

иллюстрации один соединительный элемент используют для представления множества соединений, отношений или связей между элементами. Например, если соединительный элемент представляет передачу сигналов, данных или команд, специалистам в данной области техники следует понимать, что такой элемент представляет один или множество путей сигнала, по необходимости, для воздействия на передачу.

Как указано выше, когда период кадра аудиокодека (кодера/декодера) приближается к 40 мс, или даже 20 мс или менее, основу аудиоданных можно обновлять в коротких интервалах времени. Однако в общем известно, что дополнительную информацию (или метаданные/параметр) не нужно обновлять настолько часто. Другими словами, в кодеках с короткими периодами кадра это может в общем означать, что параметры будут чрезмерно дискретизированы, если все они будут включаться в каждый кадр (как звуковой сигнал). В некоторых реализациях может существовать возможность не отправлять метаданные в каждом кадре, а только обновлять их в каждом М-ом кадре (например, вплоть до $M = 4$ в некоторых случаях). Это в общем снижало бы средний битрейт метаданных.

Учитывая вышесказанное, в широком смысле, применение техники, описанной в настоящей заявке, может применяться к любым параметрам или дополнительной информации в кодировании звука, где временная корреляция параметров превышает шаг кодека. Например (но не в качестве ограничения), процедуры временно-дифференциального энтропийного кодирования с чередованием частот могут применяться к параметрам в кодеке голосовых и звуковых служб с эффектом присутствия (IVAS), соответствующем стандарту проекта партнерства третьего поколения (3GPP), которые моделируют пространственные взаимоотношения, или в любой технологии параметрического стереофонического кодирования, которая пытается минимизировать шаг кодека ниже 40 мс. Однако, как будет понятно и очевидно специалисту в данной области, хотя варианты осуществления настоящего изобретения могут применяться к кодеку амбиофонии первого порядка (FOA) с эффектом присутствия, подход, описанный в настоящем документе, в общем применим к любому другому подходящему аудиокодеку (например, к кодекам амбиофонии более высокого порядка, HOA), где шаг или размер кадра является малым, что в общем представляло бы некоторые специфические трудности при своевременном кодировании дополнительной информации как было упомянуто выше.

Обращаясь теперь к фиг. 1, где показано схематическое изображение (упрощенной) структурной схемы кодера/декодера («кодека») 100 для кодирования и декодирования сигналов (битовых потоков) согласно одному варианту осуществления настоящего изобретения. В частности, как может быть понятно специалисту в данной области, иллюстративный пример по фиг. 1 демонстрирует кодек 100 амбиофонии первого порядка (FOA) с пространственным реконструктором (SPAR) для кодирования и декодирования битовых потоков IVAS в формате FOA. Более конкретно, как указано на фигуре, кодек 100 FOA по фиг. 1 включает как пассивное, так и активное предсказание, как будет понятно и очевидно специалисту в данной области.

В общем говоря, для кодирования кодер IVAS может содержать блок пространственного анализа и понижающего микширования, который принимает аудиоданные, включая, но без ограничения, монофонические сигналы, стереофонические сигналы, бинауральные сигналы, пространственные звуковые сигналы (например, многоканальные пространственные звуковые объекты), FOA, амбиофонию более высокого порядка (HOA) и любые другие подходящие аудиоданные. В некоторых реализациях блок пространственного анализа и понижающего микширования может

реализовывать комплексную расширенную связь (SACPL) для анализа / понижающего микширования стереофонических сигналов / звуковых сигналов FOA и/или SPAR для анализа / понижающего микширования звуковых сигналов FOA. В других реализациях блок пространственного анализа и понижающего микширования также может
5 реализовывать любые другие подходящие форматы.

Теперь снова обращаясь к фиг. 1, кодек 100 FOA может включать кодер 101 SPAR FOA, кодер 105 голосовых служб с расширенными возможностями (EVS), декодер 106 SPAR FOA и декодер 107 EVS. Кодер 101 SPAR FOA может быть выполнен с
10 возможностью преобразования входного сигнала FOA в набор каналов понижающего микширования и параметров, используемых для восстановления входного сигнала на декодере 106 SPAR FOA. В зависимости от различных реализаций, сигналы понижающего микширования могут варьировать от 1 до 4 каналов, а параметры (которые иногда также называют коэффициентами) могут содержать, но без ограничения, коэффициенты предсказания (PR), коэффициенты перекрестного предсказания (C) и коэффициенты
15 декорреляции (P). Следует отметить, что SPAR представляет собой способ, используемый для реконструкции звукового сигнала из подвергнутой понижающему микшированию версии звукового сигнала с использованием параметров PR, C и P, как будет более подробно описано ниже.

В зависимости от количества каналов понижающего микширования, один из входных
20 сигналов FOA может всегда передаваться без изменений (например, канал W, как показано в настоящем примере по фиг. 1), а от 1 до 3 других каналов (например, каналы Y, Z и X, как показано в настоящем примере по фиг. 1) могут передаваться или в виде остаточных сигналов, или полностью параметрически.

В частности, параметры предсказания могут оставаться теми же независимо от
25 количества каналов понижающего микширования и могут использоваться для сведения к минимуму предсказываемой энергии в остаточных каналах понижающего микширования. С другой стороны, параметры перекрестного предсказания могут использоваться для дополнительного содействия при восстановлении полностью параметризованных каналов из остаточных сигналов. По существу, эти параметры не
30 были бы необходимы в случаях понижающего микширования 1 и 4 каналов, где в первом случае нет остаточных каналов, из которых необходимо выполнять предсказание, а в последнем случае нет параметризованных каналов для предсказания. Кроме этого параметры декоррелятора могут использоваться для возмещения оставшейся энергии, не учтенной предсказанием и перекрестным предсказанием. Как
35 и ранее, количество параметров декорреляции может зависеть от количества каналов понижающего микширования в каждой полосе.

Пример по фиг. 1 в общем изображает иллюстративный вариант осуществления такой системы и то, как такие параметры приспособить на стороне декодера. В частности, пример реализации, показанный на фиг. 1, изображает номинальное 2-
40 канальное понижающее микширование, где представление канала W (которое представляет собой W для пассивного предсказания или W' для активного предсказания) передают на декодер 106 в неизменном виде с одним предсказанным каналом Y'. Коэффициенты (C) перекрестного предсказания позволяют восстанавливать по меньшей мере некоторую часть параметрических каналов из остаточных каналов в случаях,
45 когда по меньшей мере один канал отправлен как остаточный и по меньшей мере один отправлен параметрически, т. е. для понижающих микширований 2 и 3 каналов. Таким образом, в общем говоря, для понижающих микширований двух каналов параметры C позволяют восстанавливать некоторые из каналов X и Z из Y', а остальные каналы

восстанавливают посредством декоррелированных версий канала W , как более подробно описано ниже. В случае понижающего микширования 3 каналов остаточные каналы Y' и X' используют для восстановления только Z .

Как также будет понятно и очевидно специалисту в данной области, следует отметить, что в некоторых иллюстративных реализациях W может быть активным каналом (или, другими словами, каналом с активным предсказанием, далее называемым здесь W'). В качестве примера (но не ограничения), активный канал W , который позволяет осуществлять какой-то тип микширования каналов X , Y , Z в канал W , может быть определен следующим образом:

$$W' = W + f * pr_y * Y + f * pr_z * Z + f * pr_x * X \quad (1)$$

где f является подходящей постоянной (например, 0,5), позволяющей микшировать по меньшей мере некоторые из каналов X , Y , Z в канал W , а pr_y , pr_x и pr_z являются коэффициентами предсказания (PR). Соответственно, в случаях пассивного W , $f = 0$, поэтому микширование каналов X , Y , Z в канал W отсутствует.

В иллюстративной реализации по фиг. 1, кодер 101 SPAR FOA может содержать блок 102 (пассивного или активного) предсказателя, блок 103 повторного микширования и блок 104 выбора извлечения/понижающего микширования. В частности, предсказатель 102 может принимать каналы FOA в 4-канальном формате B (W , Y , Z , X) и вычисляет каналы понижающего микширования (представление W , Y' , Z' , X').

Блок 104 выбора извлечения/понижающего микширования может извлекать метаданные SPAR FOA, например из раздела полезной нагрузки метаданных битового потока IVAS. Блок 102 предсказателя и блок 103 повторного микширования затем могут использовать метаданные SPAR FOA для генерирования подвергнутых повторному микшированию каналов FOA (представление W , S_1' , S_2' и S_3'), которые затем могут быть поданы на кодер 105 EVS с целью их кодирования в битовый поток EVS, который может быть впоследствии инкапсулирован в битовом потоке IVAS, отправляемом на декодер 106.

Со ссылкой на декодер 106 SPAR FOA, битовый поток EVS декодируется декодером 107 EVS, что в результате дает некоторое количество (например, $N_{dmx} = 2$, где N_{dmx} обозначает количество каналов понижающего микширования) каналов понижающего микширования. В некоторых реализациях декодер 106 SPAR FOA может быть выполнен с возможностью выполнения операций, противоположных операциям, которые были выполнены кодером 101 SPAR. Например, в примере по фиг. 1 подвергнутые повторному микшированию каналы FOA (представление W , S_1' , S_2' и S_3') могут быть восстановлены из 2 каналов понижающего микширования с использованием пространственных метаданных SPAR FOA. Подвергнутые повторному микшированию каналы SPAR FOA затем могут быть введены в обратный микшер 111 для восстановления каналов понижающего микширования SPAR FOA (представление W , Y' , Z' и X'). После этого предсказанные каналы SPAR FOA могут затем быть введены в обратный предсказатель 112 для восстановления исходных, не подвергнутых микшированию каналов SPAR FOA (W , Y , Z и X).

Следует отметить, что в этом двухканальном примере для генерирования декоррелированных версий канала W с использованием декоррелятора во временной области или частотной области могут быть использованы блоки 109-1 (dec_1) и 109-2 (dec_2) декоррелятора. Каналы понижающего микширования и декоррелированные

каналы могут быть использованы в комбинации с метаданными SPAR FOA для параметрического восстановления каналов X и Z. Блок 108 С может относиться к умножению остаточного канала на матрицу коэффициентов С размером 2×1, тем самым создавая два сигнала перекрестного предсказания, которые могут быть суммированы в параметрически восстановленные каналы, как показано в примере по фиг. 1. Кроме этого, блок 110-1 P₁ и блок 110-2 P₂ могут относиться к умножению выходных сигналов декореллятора на столбцы матрицы коэффициентов Р размером 2×2, тем самым создавая четыре выходных сигнала, которые могут быть суммированы в параметрически восстановленные каналы, как показано в примере по фиг. 1.

Как указано выше, в некоторых реализациях, в зависимости от количества каналов понижающего микширования, один из входных сигналов FOA может быть передан на декодер 106 SPAR FOA без изменений (например, иллюстративный канал W), а от одного до трех других каналов (Y, Z и X) могут быть переданы на декодер 106 SPAR FOA или в качестве остаточных сигналов, или полностью параметрически.

Коэффициенты PR, которые остаются теми же вне зависимости от количества N_{dmx} каналов понижающего микширования, могут использоваться для сведения к минимуму предсказываемой энергии в остаточных каналах понижающего микширования. Коэффициенты С могут использоваться для дополнительного содействия при восстановлении полностью параметризованных каналов из остаточных сигналов. Как таковые, коэффициенты С могут не требоваться в случаях одно- и четырехканального понижающего микширования, где отсутствовали бы остаточные каналы или параметризованные каналы для предсказания на их основе. Коэффициенты Р используются для возмещения остающейся энергии, не учтенной коэффициентами PR и С. Количество коэффициентов Р обычно зависит от количества N каналов понижающего микширования в каждой полосе.

В некоторых реализациях коэффициенты PR SPAR (только пассивный W) вычисляют следующим образом.

Этап 1. Предсказать все побочные сигналы (Y, Z, X) из основного сигнала W с использованием матрицы предсказания, состоящей из коэффициентов предсказания, следующим образом:

$$\begin{bmatrix} W' \\ Y' \\ Z' \\ X' \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ -pr_Y & 1 & 0 & 0 \\ -pr_Z & 0 & 1 & 0 \\ -pr_X & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} W \\ Y \\ Z \\ X \end{bmatrix} \quad (2)$$

где, в качестве примера, параметр предсказания для предсказанного канала Y' может быть вычислен как:

$$pr_Y = \frac{R_{YW}}{\max(R_{WW}, \epsilon)} \frac{1}{\max(1, \sqrt{|R_{YY}|^2 + |R_{ZZ}|^2 + |R_{XX}|^2})} \quad (3)$$

где $R_{AB} = cov(A, B)$ являются элементами входной ковариационной матрицы, соответствующей сигналам А и В, и могут быть вычислены для каждой полосы. Подобным образом, остаточные каналы Z' и X' имеют соответствующие параметры предсказания, а именно pr_Z и pr_X . Вышеописанная матрица известна как матрица предсказания.

Этап 2. Выполнить повторное микширование W и предсказанных сигналов (Y', Z',

X') от наиболее до наименее акустически значимых, где «повторное микширование» означает изменение порядка или изменение комбинации сигналов на основе некоторой методологии,

$$\begin{bmatrix} W \\ S_1' \\ S_2' \\ S_3' \end{bmatrix} = \begin{bmatrix} \text{remix} \end{bmatrix} \begin{bmatrix} W' \\ Y' \\ Z' \\ X' \end{bmatrix} \quad (4)$$

Одна возможная реализация повторного микширования представляет собой изменение порядка входных сигналов на W, Y', X' и Z', в предположении, что звуковые метки слева и справа являются более акустически значимыми или важными, чем передне-задние, а передне-задние метки являются более акустически значимыми/важными, чем верхне-нижние метки.

Этап 3. Вычислить ковариацию 4-канального пост-предсказания и повторного микширования понижающего микширования, как:

$$R_{pr} = [\text{remix}] [\text{prediction}] \cdot R \cdot [\text{prediction}]^H [\text{remix}]^H, \quad (5)$$

где матрицы [prediction] и [remix] относятся к матрицам, используемым в уравнениях (2) и (4) соответственно. Окончательная матрица пост-предсказания и повторного микширования понижающего микширования может быть записана в виде

$$R_{pr} = \begin{pmatrix} R_{WW} & R_{Wd} & R_{Wu} \\ R_{dW} & R_{dd} & R_{du} \\ R_{uW} & R_{ud} & R_{uu} \end{pmatrix} \quad (6)$$

где d представляет остаточные каналы (т.е. каналы с 2-го по N_dmx, при этом N_dmx обозначает количество каналов понижающего микширования), а u представляет параметрические каналы, которые необходимо полностью восстановить (т.е. каналы с (N_dmx+1)-ого по 4-й).

Для примера понижающего микширования WS₁S₂S₃ с 1-4 каналами d и u могут представлять следующие каналы, показанные в таблице 1:

N	Каналы d	Каналы u
1	--	S ₁ ', S ₂ ', S ₃ '
2	S ₁ '	S ₂ ', S ₃ '
3	S ₁ ', S ₂ '	S ₃ '
4	S ₁ ', S ₂ ', S ₃ '	--

Таблица 1. Представления каналов d и u

Наибольший интерес для вычисления метаданных SPAR FOA представляют величины R_{dd}, R_{ud} и R_{uu}.

Этап 4. На основе величин R_{dd}, R_{ud} и R_{uu} кодек 100 может определить, возможно ли перекрестное предсказание какой-либо остающейся части полностью параметрических каналов из остаточных каналов, передаваемых на декодер. В некоторых возможных реализациях необходимые дополнительные коэффициенты C можно вычислить как:

$$C = R_{dd}(R_{dd} + I \max(\epsilon, \text{tr}(R_{dd}) * 0.005))^{-1}. \quad (7)$$

Следовательно, параметр C в общем будет иметь форму (1×2) для 3-канального понижающего микширования и (2×1) для 2-канального понижающего микширования.

Этап 5. Вычислить остающуюся энергию в параметризованных каналах, которые необходимо восстановить декорреляторами 109-1 и 109-2, как:

$$Reg_{uu} = CR_{dd}C^H \quad (8)$$

$$Res_{uu} = R_{uu} - Reg_{uu} \quad (9)$$

$$P = \sqrt{\frac{Res_{uu}}{\max(\epsilon, R_{WW}, \alpha * \text{tr}(|Res_{uu}|))}} \quad (10)$$

где $0 \leq \alpha \leq 1$ является постоянным коэффициентом масштабирования. Следует

отметить, что остаточная энергия в каналах Res_{uu} повышающего микширования представляет собой разность между фактической энергией R_{uu} (пост-предсказание) и восстановленной энергией Reg_{uu} перекрестного предсказания.

В некоторых возможных реализациях, после того, как недиагональным элементам нормализованной матрицы Res_{uu} были присвоены нулевые значения, из матрицы можно извлечь квадратный корень. P также может представлять собой ковариационную матрицу, а значит являться эрмитово-симметричной. Таким образом, на декодер 106 необходимо передавать только параметры из верхнего или нижнего треугольника. Диагональные элементы могут являться вещественными, тогда как недиагональные элементы могут быть комплексными. В некоторых дополнительных возможных реализациях коэффициенты P могут быть дополнительно разделены на диагональные и недиагональные элементы P_d и P_o , соответственно. В некоторых реализациях только диагональные элементы коэффициентов P вычисляют и передают на декодер, и их можно вычислить следующим образом:

$$P_d = \sqrt{\frac{\text{diag}(Res_{uu})}{\max(\epsilon, R_{WW}, \alpha * \text{tr}(|Res_{uu}|))}}$$

Теперь, на стороне кодера, квантование этих параметров может стать необходимым. В частности, учитывая зависимости между тремя типами параметров (т.е. PR , C и P), как указано выше, порядок (или последовательность) их вычисления и квантования может в общем считаться важным для качества звучания. Согласно настоящему изобретению, ниже представлены три возможных варианта осуществления способов достижения этого:

1. Все в одном

В этом варианте осуществления декорреляторам в общем запрещено компенсировать погрешности квантованного предсказания.

Точнее, на первом этапе вычисляют параметры PR , затем C , а затем P , как изображено выше, без квантования. Затем все параметры PR , C и P квантуют согласно стратегии

или схеме квантования (например, на основе подходящих диапазонов квантования и/или уровней квантования, как будет понятно специалисту в данной области).

2. Каскадный

В общем говоря, этот конкретный вариант осуществления позволяет выполнять точное предсказание и перекрестное предсказание, а декорреляторы могут возмещать погрешности, вызванные квантованием.

Точнее, на первом этапе вычисляют и затем квантуют параметр PR. После этого из квантованных параметров PR вычисляют и затем квантуют параметр C. Наконец, из квантованных параметров C также вычисляют и затем квантуют параметр P.

3. Частичный каскадный

В общем говоря, этот конкретный вариант осуществления сводил бы к минимуму коэффициенты P, тем самым позволяя выполнять точное перекрестное предсказание, но не позволяя декорреляторам компенсировать погрешности предсказания.

Точнее, на первом этапе параметры PR, C и P вычисляют без квантования, как в вышеописанном варианте осуществления «все в одном», а затем квантуют параметр P. После этого также квантуют параметры PR. И наконец, из квантованных параметров PR повторно вычисляют и затем квантуют параметр C.

В каждом из изображенных выше вариантов осуществления понижающее микширование (включая остаточные сигналы) всегда можно вычислить с использованием квантованных коэффициентов предсказания.

Как будет понятно и очевидно специалисту в данной области, собственно процесс квантования можно определить с помощью подходящего диапазона (квантования). Например, диапазон [-a, a] может быть определен для некоторых параметров (например, параметров PR, C и недиагональных элементов параметра P), в то время как другой диапазон [0, a] может быть определен для других. Кроме этого, также могут быть определены несколько уровней квантования, которые должны быть равномерно распределены между этими конечными точками. Другими словами, различные ограничения и величины шага могут быть настроены или определены для каждого типа параметра (например, PR, C, P_d, P_o). Кроме этого, в некоторых реализациях, если параметры являются комплексными значениями, действительные и мнимые части могут быть квантованы с использованием одинаковых/разных диапазонов и количества шагов, согласно распределению параметров.

Возможная реализация процесса квантования может быть определена как:

$$q(x) = \max(-a, \min(a, x)) / (2a / (qlvl - 1)) \quad (11)$$

или

$$q(x) = \max(0, \min(a, x)) / (a / (qlvl - 1)) \quad (12)$$

где x обозначает индексы квантования, a обозначает диапазон квантования и qlvl обозначает уровень квантования.

В некоторых возможных реализациях может быть желательно выбрать нечетные значения для уровней квантования (т.е. qlvl), чтобы обеспечить доступность точки квантования в 0, например для двусторонних параметров, как будет очевидно специалисту в данной области.

Будет смысл отметить, как уже было указано выше, что пример по фиг. 1 в общем показывает реализацию пассивного предсказания (т.е. канала W). Однако, как будет понятно и очевидно специалисту в данной области, в некоторых других возможных

реализациях может применяться активное предсказание. В общем говоря, активный канал W может позволять выполнять некоторый вид микширования по меньшей мере некоторых каналов X, Y, Z в канал W, и такое активное предсказание обычно может использоваться в случае 1-канального понижающего микширования. Соответственно, в случаях пассивного предсказания обычно будет отсутствовать микширование каналов X, Y, Z в канал W.

Фиг. 2 представляет собой блок-схему, изображающую пример способа 200 покадрового кодирования метаданных для входного сигнала согласно одному варианту осуществления настоящего изобретения. Способ 200, как описано в настоящем документе, может быть применен, например, к кодеку 100, который показан на фиг. 1 (или к любому другому подходящему кодеку). Метаданные можно рассчитать/вычислить (например, извлечь) из входного (звукового или видео) сигнала путем использования подходящего кодека (кодера/декодера). В общем говоря, метаданные можно использовать для оказания помощи в восстановлении входного сигнала на стороне декодера. Метаданные могут содержать множество по меньшей мере частично взаимосвязанных параметров, которые могут быть вычислены из входного сигнала. Другими словами, по меньшей мере некоторые параметры входного сигнала могут быть вычислены (например, сгенерированы или восстановлены) в зависимости от по меньшей мере некоторых других параметров, так что, в зависимости от различных обстоятельств, не все параметры необходимо всегда передавать явным образом.

Способ 200 можно выполнять итеративно, например с использованием зацикленного процесса (который будет подробно описан ниже) для каждого кадра входного сигнала. В частности, способ 200 (или точнее, зацикленный процесс) начинается этапом S210, на котором определяют стратегию обработки из множества стратегий обработки для вычисления и квантования параметров.

После определения (например, выбора) стратегии обработки на этапе S210, зацикленный процесс переходит к этапу S220, на котором выполняют вычисление и квантование параметров на основе определенной стратегии обработки, чтобы получить квантованные параметры.

После этого на этапе S230 (квантованные) параметры соответственно кодируют, и затем (итоговый) битрейт оценивают (например, вычисляют) из закодированных параметров и принимают решение на основе оцененного битрейта вместе с по меньшей мере одним целевым пороговым значением битрейта (например, предварительно определенным или предварительно сконфигурированным) на этапе S240.

Если условие порогового значения битрейта выполнено, например оцененный битрейт равен пороговому значению битрейта или меньше его, способ 200 выходит из цикла обработки. Иначе цикл возвращается на этап S210 и продолжается этапами S210-S240. В частности, при повторном входе в цикл может быть определена новая стратегия обработки, чтобы удовлетворить условие целевого порогового значения битрейта.

Как будет понятно и очевидно специалисту в данной области, множество стратегий обработки для вычисления и квантования параметров могут быть предоставлены любым подходящим образом, например предварительно определенным или предварительно сконфигурированным. Соответственно, стратегия обработки также может быть определена из множества стратегий обработки любым подходящим образом. Например, в зависимости от (текущего) требования к битрейту, подходящая стратегия обработки может быть выбрана из множества стратегий обработки, так что итоговый битрейт после выполнения вычисления, квантования и кодирования (например, с энтропийным кодированием или без него) на основе выбранной таким образом стратегии

обработки удовлетворяет (текущему) требованию к битрейту.

Поскольку зацикленный процесс в общем направлен (помимо прочего) на обработку, связанную с квантованием, в некоторых случаях зацикленный процесс также может называться циклом квантования (или просто циклом в целях краткости). Подобным образом, поскольку стратегия обработки также в общем направлена (помимо прочего) на обработку, связанную с квантованием, в некоторых случаях стратегия обработки также может называться стратегией квантования (или, в некоторых других случаях, взаимозаменяемо называться схемой квантования). Кроме этого необходимо отметить, что процесс кодирования может использовать любую подходящую процедуру кодирования, включающую, но без ограничения, энтропийное кодирование или кодирование без энтропии (например, кодирование с основанием 2). Разумеется, в зависимости от различных реализаций и/или требований может использоваться любой другой подходящий механизм кодирования.

В частности, каждая из множества стратегий обработки может содержать соответствующий первый указатель, указывающий на порядок (или последовательность) применительно к вычислению и квантованию отдельных параметров. Другими словами, первый указатель может содержать информацию о последовательности, указывающую, когда и в каком порядке выполнять вычисление и квантование отдельных параметров. В качестве примера (но не ограничения), первый указатель может содержать информацию, указывающую, что все параметры сначала вычисляют, перед тем как выполнять квантование любого из них.

Далее зацикленный процесс будет описан более подробно со ссылкой на примеры, показанные на фиг. 3 и 4.

Как указано выше, в кодеках с короткими шагами или обновлениями кадра параметры могут быть чрезмерно дискретизированы, если все они включаются в каждый кадр. Таким образом, главная цель настоящего изобретения заключается в предоставлении механизмов для того, чтобы насколько возможно свести к минимуму дополнительную информацию, но в то же время сохранить высокую частоту обновления кадров для основы аудиоданных и параметров.

В широком смысле, для решения вышеописанной задачи, в частности для оценки распространения дополнительной информации, автор настоящего изобретения в общем предлагает механизм внедрения временно-дифференциальных оценок для параметров некоторых (частотных) полос вместе с недифференциальными оценками для параметров других (частотных) полос. Предложенный подход чередует кодирование полос временно-дифференциальным образом и недифференциальным образом так, что каждую полосу регулярно обновляют недифференциальным вычислением без необходимости в полном обновлении параметров. Ключевая идея заключается в том, что по мере уменьшения размера кадра корреляция параметров от кадра к кадру увеличивается и, таким образом, с помощью временно-дифференциального кодирования параметров можно добиться повышенной эффективности кодирования.

В дополнение к частотному чередованию временно-дифференциального кодирования также внедрена идея итеративного и поэтапного подхода к выбору оптимальной схемы квантования параметров, который ищет «наилучшую» (или оптимальную) схему квантования из нескольких альтернатив. В этом случае термин «лучшая» или «оптимальная» не обязательно может относиться к схеме квантования с наименьшим битрейтом параметров, а к схеме, которая облегчает состояние для декодера.

Например, использование временно-дифференциального кодирования в общем может обладать недостатком, который в первую очередь заключается в том, что

вводится состояние от кадра к кадру, которое может стать причиной проблем, когда во время передачи звуковой поток может столкнуться с потерей пакетов. В этом случае как звуковые данные, так и параметры могут быть утеряны и любые параметры, обновляемые с помощью временно-дифференциального кодирования, могут столкнуться с потенциальными артефактами в нескольких последующих кадрах. В настоящем изобретении в общем не рассматривается уменьшение указанной проблемы с помощью декодера. Вместо этого данную проблему в общем решают (облегчают) путем выбора подходящей схемы квантования, которая будет насколько возможно ограничивать такое поведение. В широком смысле уменьшение с помощью кодирования (стороной кодера) в общем включает итеративный процесс выбора квантования и энтропийного кодирования, который пытается свести к минимуму возможный масштаб появления артефактов из-за потери пакетов в результате использования временно-дифференциального кодирования.

Теперь обращаясь назад к фигурам, фиг. 3 представляет собой блок-схему, схематично изображающую пример цикла 300 обработки согласно одному варианту осуществления настоящего изобретения.

Цикл 300 обработки начинается этапом S310, на котором вычисляют (или оценивают) первый битрейт (здесь и далее обозначаемый b_1). В некоторых возможных реализациях для каждого кадра оценивают энтропию недифференциально и/или частотно-дифференциально квантованных параметров. В некоторых других возможных реализациях первый битрейт b_1 можно вычислить как минимум среди схем недифференциального и частотно-дифференциального кодирования, кодируемых с использованием (обученных) энтропийных кодеров (например, кодирование Хаффмана или арифметическое кодирование).

На этапе S320 первый битрейт b_1 сравнивают с целевым битрейтом (здесь и далее обозначаемым t). Если оцененный битрейт b_1 параметров находится в рамках (то есть равен или меньше) целевого битрейта t , то выполняют выход из цикла обработки. В результате параметры кодируют так, что любые дополнительные доступные биты предоставляют на аудиокодер для увеличения битрейта основы аудиоданных.

Если этап S320 завершается неудачей (т.е. оцененный битрейт b_1 больше целевого битрейта t), то на этапе S330 вычисляют второй битрейт (здесь и далее обозначаемый b_2) квантованных параметров. В некоторых возможных реализациях второй битрейт b_2 можно вычислить недифференциальным образом без энтропийного кодирования (например, с использованием кодирования с основанием 2).

Затем на этапе S340 второй битрейт b_2 сравнивают с целевым битрейтом t . Если второй битрейт b_2 находится в рамках (то есть равен или меньше) целевого битрейта t , то выполняют выход из цикла обработки.

Иначе на этапе S350 вычисляют третий битрейт (здесь и далее обозначаемый b_3) параметров. В некоторых возможных реализациях третий битрейт b_3 можно вычислить временно-дифференциальным кодированием с использованием (обученных) энтропийных кодеров. В некоторых дополнительных возможных реализациях подгруппу значений параметров в текущем кадре можно квантовать, а затем вычесть из квантованных значений параметров в предыдущем кадре, и можно вычислить дифференциальное квантованное значение и энтропию параметра.

На этапе S360, если вычисленный битрейт b_3 равен пороговому значению t или меньше него, то выполняют выход из цикла обработки и параметры кодируют с предоставленным битрейтом и дополнительные биты предоставляют для кодирования звуковых данных.

Иначе на этапе S370 могут быть реализованы различные меры для итогового обеспечения соответствия целевому пороговому значению битрейта t .

Например, в некоторых возможных реализациях вторая, более грубая, стратегия обработки (стратегия квантования) может быть выбрана из множества стратегий обработки. В таких случаях, как будет понятно и очевидно специалисту в данной области, процесс квантования может включать несколько уровней все более грубого квантования, таких как, например, стратегии точного, среднего, грубого и очень грубого квантования. Затем, после определения (например, выбора) более грубой стратегии квантования, цикл обработки повторяет этапы с S310 по S360.

В некоторых других возможных реализациях уменьшение количества полос частот может быть выполнено на этапе S370. Затем вышеупомянутые этапы (т.е. этапы с S310 по S360) можно повторить с уменьшенной конфигурацией полос. Это в общем уменьшит общее количество параметров для квантования и часто может дать в результате низкий битрейт для (по меньшей мере) некоторых кадров.

Альтернативно или дополнительно еще в некоторых других реализациях также можно выполнять этап заморозки (т.е. повторного использования) параметров в полосе из предыдущего кадра. Это по сути предотвратит изменение параметра с течением времени, что приведет к уменьшению энтропии для временно-дифференциального энтропийного кодирования. Например, как изображено в таблице 2 (которая будет подробно описана ниже), при кодировании с использованием схемы 4а кодирования можно заморозить параметры в частотных полосах 2, 6, и 10. Это, как правило, приводит к уменьшению энтропии, отсутствию изменения декодера или схемы энтропийного кодирования и к небольшому влиянию на качество. Необходимо отметить, что вышеуказанный пример с 2, 6 и 10 является лишь иллюстративным примером и много конфигураций полос могут быть заморожены в нескольких кадрах, как будет понятно и очевидно специалисту в данной области. Например, если заморозить все полосы частот в периоде из 2 кадров, то кодер может отправлять половину полос в кадре N и остальную половину в кадре $N+1$ (тем самым уменьшая общее количество параметров для отправки), что в общем означает, что декодер будет получать все (например, 12) обновленных полос частот раз в два кадра. В таких случаях, если один кадр утерян, обычно остается возможность экстраполяции из двух последних пригодных кадров. При восстановлении после потери пакетов можно выполнить интерполирование между полосами, которые были приняты с заданным кадром.

Следует отметить, что если выход из цикла выполняют на этапе x , то окончательный битрейт параметров представляет собой битрейт, вычисленный на этом этапе x .

Кроме этого, в некоторых реализациях можно (или даже желательно) рассмотреть проектирование битрейта $b3$ с самой грубой стратегией квантования (из заданного множества стратегий квантования, доступных для квантования параметров) как гарантирующее битрейт, который будет меньше целевого порогового значения t битрейта. В таких случаях можно гарантировать постоянное наличие решения для подходящего битрейта параметров в рамках целевого битрейта t .

Фиг. 4 представляет собой блок-схему, изображающую пример цикла 400 обработки согласно другому варианту осуществления настоящего изобретения. В частности, идентичные или подобные ссылочные номера в цикле 400 по фиг. 4 в общем указывают на идентичные или подобные элементы в цикле 300, которые изображены на фиг. 3, так что их повторное описание может быть пропущено в целях краткости.

В частности, цикл обработки по фиг. 4 может особенно подходить для случаев, где используются два пороговых значения битрейта (представленных как целевое пороговое

значение t_1 битрейта и максимальное пороговое значение t_2 битрейта), как противоположность сценарию с одним целевым пороговым значением битрейта, изображенным на фиг. 3. В широком смысле, целевое пороговое значение t или t_1 битрейта может считаться целью или заданием, которого желательно достичь, в то время как максимальное пороговое значение t_2 битрейта может рассматриваться просто как «жесткое» пороговое значение, которое не следует превышать.

Более конкретно, этапы с S410 по S470 подобны этапам (т.е. этапам с S310 по S370), представленным на фиг. 3, так что их повторное описание можно пропустить в целях краткости.

Однако вместо непосредственного переключения на этап S470, если условие этапа S460 не выполнено, вводят дополнительный этап S461 с вычислением четвертого битрейта (b_4) как минимума среди битрейтов b_1 , b_2 и b_3 . Затем четвертый битрейт b_4 сравнивают с максимальным пороговым значением t_2 битрейта на этапе S462.

Если четвертый битрейт b_4 равен максимальному пороговому значению t_2 битрейта или меньше него, выполняют выход из цикла 400 обработки; иначе цикл 400 обработки продолжается этапом S470 (который по сути подобен этапу S370 по фиг. 4) и повторяют этапы с S410 по S462.

Подобно изображенному на фиг. 3, если выход из цикла выполняют на этапе x , то окончательный битрейт параметров представляет собой битрейт, вычисленный на этом этапе x .

Более того, в некоторых реализациях также можно (или даже желательно) рассмотреть проектирование битрейта b_3 с самой грубой стратегией квантования (из заданного множества стратегий квантования, доступных для квантования параметров) как гарантирующее битрейт, который будет меньше максимального порогового значения t_2 битрейта. В таких случаях можно гарантировать постоянное наличие решения для подходящего битрейта параметров в рамках максимального битрейта t_2 .

Подводя итоги, этапы S310, S330 и S350 по фиг. 3 и, соответственно, также этапы S410, S430 и S450 по фиг. 4 в общем не влияют на качество звучания. Однако этап S461 по фиг. 4 мог бы ухудшить качество из-за влияния как на битрейт звуковых данных, так и на битрейт параметров. Кроме того, любые возможные технологии, упомянутые выше, на этапах S370 по фиг. 3 и S470 по фиг. 4 (например, переход к более грубому квантованию, сокращение полосы путем уменьшения разрешающей способности по частоте, сокращение полосы путем уменьшения разрешающей способности по времени и т.д.) в основном будут оказывать отрицательное влияние на качество. Таким образом, этапы в примерах по фиг. 3 и 4 расположены в таком порядке, чтобы свести к минимуму ухудшение качества или чтобы справиться с ограничениями в других областях. В широком смысле способ, описанный в настоящем изобретении, стремится выбирать одну или более из изображенных выше технологий для сохранения баланса между уменьшением битрейта метаданных и воспринимаемым качеством.

Существуют также два дополнительных рассуждения относительно конкретного порядка вышеописанных этапов и причины возможного наличия двух целевых битрейтов параметров (т.е. битрейтов t_1 и t_2).

В частности, поэтапный порядок позволяет прервать процедуру, если удовлетворены ограничения. Это в общем уменьшит вычислительную нагрузку при последовательном выполнении вычислений, поскольку обычно не нужно будет проходить все доступные этапы.

Кроме этого, упорядочивание также обеспечивает неявное наличие альтернатив. Например, расположение недифференциального энтропийного кодирования в качестве

первого этапа обычно будет означать, что эта альтернатива является предпочтительной, если она удовлетворяет ограничениям. Это представляет собой упрощение кодера для минимизации состояния с целью улучшения качества в условиях потери пакетов.

5 Более того, возможность использования двух целей (t_1 и t_2) в общем предоставляет больше возможности для управления компромиссом между битрейтом звуковых данных и битрейтом параметров.

Далее будет подробнее описано чередование для достижения временно-дифференциального кодирования.

10 Некоторые возможные реализации для управления чередованием временно-дифференциального энтропийного кодирования показаны в таблице 2.

Схема кодирования	Временно-дифференциальное кодирование, полосы 1-12
основание	000000000000
4a	011101110111
4b	101110111011
15 4c	110111011101
4d	111011101110

Таблица 2. Чередующиеся схемы временно-дифференциального кодирования

В этом конкретном примере в общем предложено 5 конфигураций для кодирования битового потока метаданных, каждая из которых состоит из 12 (частотных) полос. 20 Более конкретно, полосу, обозначенную цифрой 0, кодируют недифференциально, а полосу, обозначенную цифрой 1, кодируют временно-дифференциально (т.е. квантуют параметр и вычитают из квантованного параметра в предыдущем кадре).

Как описано в этом примере, битрейт параметров каждого кадра сначала оценивают путем кодирования недифференциальным образом (т.е. основания) с помощью 25 квантования параметров (например, см. этап S410 или S510). Затем, на этапе S450 или S550 схему временно-дифференциального кодирования выбирают (при необходимости) на основе схемы кодирования предыдущего кадра.

30 Пример отображения из схемы кодирования предыдущего кадра в схему временно-дифференциального кодирования текущего кадра показан ниже в таблице 3:

Схема кодирования предыдущего кадра	Схема временно-дифференциального кодирования текущего кадра
основание	4a
4a	4b
4b	4c
35 4c	4d
4d	4a

Таблица 3. Отображение схем временно-дифференциального кодирования

Следует отметить, что в настоящем примере термин «основание», используемый в 40 таблице 3, в общем относится к схеме недифференциального кодирования. Таким образом, как видно из таблицы 3, временно-дифференциальное кодирование всегда циклически проходит от 4a до 4d (и обратно). Возможно продолжать циклическое прохождение, даже не требуя реализации недифференциального кодирования. И в этом конкретном примере максимальной запоминаемой информацией или «состоянием» 45 кодека является текущий кадр и три предыдущих кадра (т.е. всего четыре кадра). Разумеется, как будет понятно и очевидно специалисту в данной области, такие количества как 5 конфигураций и 12 (частотных) полос и т.д. используются лишь в качестве примеров с целью иллюстрации, и может использоваться любое другое подходящее количество, в зависимости от различных реализаций и/или требований.

Аналогичные или подобные аргументы применяются к переключению между схемами кодирования, как изображено в таблице 3, которое может подобным образом использовать любую подходящую методику.

Следует отметить, если выбрана другая схема квантования, то индексы из предыдущего кадра, квантованные с использованием другой схемы квантования, сначала могут быть отображены в индексы текущего кадра. В общем говоря, этап отображения может быть необходим для обеспечения возможности временно-дифференциального кодирования параметров, например когда количество уровней квантования изменяется от одного кадра к следующему, тем самым позволяя выполнять временно-дифференциальное кодирование между кадрами, не прибегая к передаче недифференциального кадра при каждом изменении схемы квантования.

В качестве возможного примера, отображение индексов может быть выполнено на основе формул:

$$index_{cur} = round(index_{prev} \times (quant_lvl_{cur} - 1) / (quant_lvl_{prev} - 1)) \quad (13)$$

где $index_{cur}$ обозначает индексы текущего кадра после отображения, $index_{prev}$ обозначает индексы предыдущего кадра, $quant_lvl_{cur}$ обозначает уровень квантования текущего кадра и $quant_lvl_{prev}$ обозначает уровень квантования предыдущего кадра.

В качестве простого иллюстративного примера предположим, что диапазон квантования составляет от 0 до 2, и предположим, что предыдущие уровни квантования составляют 11. В случае равномерного квантования это в общем означало бы, что каждый этап квантования составлял бы 0,2. Далее предположим, что текущие уровни квантования составляют 21, что означает, что каждый этап квантования составляет 0,1 при равномерном квантовании. На основе этих предположений, если квантованное значение в предыдущем кадре составляло 0,4, то при 11 равномерных уровнях

квантования получают следующий предыдущий индекс $index_{prev} = 2$. Отображение предоставляет квантованные индексы метаданных предыдущего кадра как если бы они были квантованы с использованием уровней квантования текущего кадра. Таким образом, в этом примере, если уровни квантования в текущем кадре составляют 21, то квантованное значение 0,4 будет отображено в $index_{curr} = 4$. После того, как

отображенные индексы вычислены, вычисляют разницу между индексами текущего кадра и предыдущего кадра и кодируют эту разницу. Аналогичные или подобные подходы при необходимости также могут применяться к частотно-дифференциальному кодированию, как будет понятно и очевидно специалисту в данной области.

Разумеется, могут использоваться любые другие подходящие схемы отображения (например, с использованием справочной таблицы или подобного), в зависимости от различных реализаций и/или требований.

Более того, как указано выше, один параметр метаданных можно квантовать из непрерывного числового значения в индекс, представляющий собой дискретное значение. При недифференциальном кодировании информация, кодируемая для этого параметра метаданных, соответствует непосредственно этому индексу. При временно-дифференциальном кодировании кодируемая информация представляет собой разницу

между индексом параметра метаданных из текущего кадра и индексом того же параметра метаданных из предыдущего кадра. Как будет понятно и очевидно специалисту в данной области, вышеописанная общая идея временно-дифференциального кодирования может быть дополнительно расширена, например на множество полос частот. Соответственно, параметр метаданных может быть расширен подобным образом, например на множество параметров, соответственно относящихся ко множеству полос частот, при необходимости. Частотно-дифференциальное кодирование следует подобному принципу, но закодированная разница является разницей между метаданными одной полосы частот текущего кадра и метаданными другой полосы частот текущего кадра (в отличие от текущего кадра минус предыдущий кадр во временно-дифференциальном кодировании). В качестве простого примера (но не ограничения), предположим, что a_0 , a_1 , a_2 и a_3 обозначают индексы параметров в 4 полосах частот конкретного кадра, тогда в одном примере реализации частотно-дифференциальные индексы могут представлять собой a_0 , a_0-a_1 , a_1-a_2 , a_2-a_3 . Как будет очевидно специалисту в данной области, общая идея (временно-и/или частотно-) дифференциального кодирования заключается в том, что метаданные обычно могут медленно изменяться от кадра к кадру или от полосы частот к полосе частот, так что даже если исходное значение метаданных было большим, разница между ним и метаданными предыдущего кадра или разница между ним и метаданными другой полосы частот вероятно будет небольшой. Это является преимущественным, поскольку в общем параметры со статистическими распределениями, которые стремятся к нулю, могут быть закодированы с использованием меньшего количества битов. Таким образом, даже если некоторые из иллюстративных реализаций кратко или просто ссылаются на временно-дифференциальное кодирование, специалисту в данной области будет понятно, что также к ним может применяться и частотно-дифференциальное кодирование (возможно с небольшой подходящей адаптацией).

Некоторые дополнительные возможные примеры настоящего изобретения могут относиться к процессу обработки входного звукового сигнала, представленного в подполосах, для получения подвергнутого понижающему микшированию сигнала и связанных с ним метаданных, который может быть выполнен одним или более процессорами. Процесс может включать, для каждой подполосы, определение матрицы понижающего микширования и связанных метаданных; и повторное микширование каждой из указанных подполос согласно указанной матрице понижающего микширования для получения указанного подвергнутого понижающему микшированию сигнала. Одна или более стратегий квантования и одна или более стратегий кодирования могут использоваться для кодирования метаданных, учитывая ограничение целевого и/или максимального битрейта метаданных.

В некоторых реализациях процесс может включать недифференциальное энтропийное кодирование всех подполос. Процесс может дополнительно включать частотно-дифференциальное энтропийное кодирование всех подполос. Процесс может дополнительно включать комбинирование частотного чередования с временно-дифференциальным кодированием квантованных параметров, соответствующих выбранным подполосам, для аудиокодека с низкой задержкой, как подробно описано ниже.

Процесс может дополнительно включать не являющееся энтропийным кодирование метаданных подполос. Итеративное прохождение этапов для поиска подходящей стратегии кодирования для выполнения требований к битрейту и качеству звучания и уменьшения состояния декодера. Процесс может дополнительно включать уменьшение

разрешающей способности по частоте путем уменьшения количества подполос, в которых будут кодированы пространственные метаданные, например с 12 полос до 6 полос. Процесс может включать уменьшение разрешающей способности по времени путем фиксации во времени (или замораживания) метаданных одной или более подполос, так что метаданные подполос не нужно передавать. Процесс может включать использование нескольких стратегий квантования, где каждая стратегия представляет собой комбинацию уровней квантования для различных параметров пространственных метаданных, причем процесс может дополнительно включать выбор одной из этих трех стратегий квантования для обеспечения достижения целевых битрейтов. Процесс может включать итеративное выполнение этапов для нахождения подходящей схемы квантования для выполнения требований к битрейту и качеству звучания. Итеративный способ направлен на получение желаемого битрейта метаданных с желаемой схемой квантования, минимальной вычислительной сложностью и уменьшенным состоянием декодера. Если желаемый уровень квантования не попадает в желаемый диапазон битрейтов, то возвращаются к (например, более грубой) схеме квантования путем обеспечения минимального влияния на качество звучания.

В некоторых реализациях отображение индексов из предыдущих кадров, квантованных в другое количество уровней, в индексы текущего кадра позволяет выполнять временно-дифференциальное кодирование между кадрами, не прибегая к передаче недифференциального кадра каждый раз, когда необходим другой уровень квантования.

В различных реализациях квантование (преобразование непрерывных значений в дискретные индексы для кодирования) может включать определение наилучшего значения для коэффициентов согласно текущим потребностям, путем изменения порядка вычисления и квантования последовательных коэффициентов метаданных.

Вычислительное устройство, реализующее описанные выше методики, может иметь следующую иллюстративную архитектуру. Возможны и другие архитектуры, включая архитектуры с большим или меньшим количеством компонентов. В некоторых реализациях иллюстративная архитектура включает один или более процессоров (например, двухъядерные процессоры Intel® Xeon®), одно или более устройств вывода (например, ЖК-дисплей), один или более сетевых интерфейсов, одно или более устройств ввода (например, мышь, клавиатуру, сенсорный дисплей) и один или более машиночитаемых носителей (например, RAM, ROM, SDRAM, жесткий диск, оптический диск, флеш-память и т.д.). Эти компоненты могут обмениваться сообщениями и данными по одному или более каналам связи (например, шинам), которые могут использовать различное аппаратное обеспечение и программное обеспечение для облегчения передачи данных и управляющих сигналов между компонентами.

Термин «машиночитаемый носитель» относится к носителю, который участвует в предоставлении команд на процессор для исполнения, включая, но без ограничения, энергонезависимые носители (например, оптические или магнитные диски), энергозависимые носители (например, память) и среды передачи. Среда передачи включает, без ограничения, коаксиальные кабели, медный провод и оптоволокно.

Машиночитаемый носитель может дополнительно включать операционную систему (например, операционную систему Linux®), модуль передачи данных по сети, диспетчер звуковых интерфейсов, диспетчер обработки звука и распределитель содержимого в прямом эфире. Операционная система может быть многопользовательской, многопроцессорной, многозадачной, многопоточной, работать в реальном времени и т.д. Операционная система выполняет базовые задачи, включая, но без ограничения:

распознавание входящих данных с сетевых интерфейсов 706 и/или устройств 708 и передачу на них выходных данных; отслеживание и управление файлами и директориями на машиночитаемых носителях (например, в памяти или запоминающем устройстве); управление периферийными устройствами; и управление трафиком одного или более каналов связи. Модуль передачи данных по сети включает различные компоненты для 5 установления и поддержания сетевых соединений (например, программное обеспечение для реализации протоколов связи, таких как TCP/IP, HTTP и т.д.).

Архитектура может быть реализована в инфраструктуре параллельной обработки или одноранговой инфраструктуре или на одном устройстве с одним или более 10 процессорами. Программное обеспечение может включать несколько программных компонентов или может представлять собой одно тело кода.

Описанные признаки могут быть преимущественным образом реализованы в одной или более компьютерных программах, которые могут быть исполнены на программируемой системе, включающей по меньшей мере один программируемый 15 процессор, соединенный с возможностью приема данных и команд из системы хранения данных и их передачи в нее, по меньшей мере одно устройство ввода и по меньшей мере одно устройство вывода. Компьютерная программа представляет собой набор команд, которые могут использоваться, прямо или опосредованно, в компьютере для выполнения определенной деятельности или для получения определенного результата. 20 Компьютерная программа может быть написана на языке программирования любой формы (например, Objective-C, Java), включая компилируемые или интерпретируемые языки, и она может быть развернута в любой форме, например как самостоятельная программа или как модуль, компонент, подпрограмма, веб-приложение, выполняемое в браузере, или другой элемент, подходящий для использования в вычислительной 25 среде.

Подходящие процессоры для исполнения программы, состоящей из команд, включают, в качестве примера, микропроцессоры как общего, так и специального назначения, и единственный процессор или один из нескольких процессоров или ядер в компьютере любого типа. В целом, процессор будет принимать команды и данные 30 из постоянного запоминающего устройства или оперативного запоминающего устройства, или из обоих. Ключевыми элементами компьютера являются процессор, предназначенный для исполнения команд, и одно или более устройств памяти, предназначенных для хранения команд и данных. В общем, компьютер также может содержать, или может быть функционально соединен для обмена данными, одно или 35 более запоминающих устройств большой емкости для хранения файлов данных; такие устройства включают магнитные диски, такие как внутренние жесткие диски и съемные диски; магнитооптические диски; и оптические диски. Запоминающие устройства, подходящие для материального воплощения компьютерных программных команд и данных, включают все формы энергонезависимой памяти, включая в качестве примера 40 полупроводниковые устройства памяти, такие как EPROM, EEPROM и устройства флеш-памяти; магнитные диски, такие как внутренние жесткие диски и съемные диски; магнитооптические диски; и диски CD-ROM и DVD-ROM. Процессор и память могут быть дополнены интегральными схемами специального назначения (ASIC) или могут быть интегрированы в такие схемы.

Для обеспечения взаимодействия с пользователем признаки могут быть реализованы на компьютере, содержащем устройство отображения, такое как монитор CRT (с электронно-лучевой трубкой) или LCD (с жидкокристаллическим дисплеем) или устройство с дисплеем Retina, для отображения информации пользователю. Компьютер

может содержать устройство ввода с сенсорной поверхностью (например, сенсорный экран) или клавиатуру и координатно-указательное устройство, такое как мышь или шаровой манипулятор, посредством которого пользователь может предоставить входные данные в компьютер. Компьютер может содержать устройство голосового ввода для приема голосовых команд от пользователя.

Признаки могут быть реализованы в компьютерной системе, содержащей серверный компонент, такой как сервер данных, или содержащей компонент межплатформенного программного обеспечения, такой как сервер приложений или интернет-сервер, или содержащей клиентский компонент, такой как клиентский компьютер, содержащий графический интерфейс пользователя или интернет-браузер, или любую их комбинацию. Компоненты системы могут быть соединены посредством любой формы или среды цифровой передачи данных, такой как сеть связи. Примеры сетей связи включают, например, LAN, WAN, а также компьютеры и сети, образующие Интернет.

Вычислительная система может включать клиенты и сервера. Клиент и сервер обычно удалены друг от друга и, как правило, взаимодействуют через сеть связи. Взаимосвязь клиента и сервера возникает посредством компьютерных программ, работающих на соответствующих компьютерах и имеющих друг с другом взаимосвязь клиент-сервер. В некоторых вариантах осуществления сервер передает данные (например, страницу HTML) на клиентское устройство (например, с целью отображения данных пользователю и приема пользовательского ввода от пользователя, взаимодействующего с клиентским устройством). Данные, генерируемые на клиентском устройстве (например, в результате взаимодействия с пользователем), могут быть приняты с клиентского устройства на сервере.

Система из одного или более компьютеров может быть сконфигурирована для выполнения конкретных действий благодаря наличию программного обеспечения, программно-аппаратного обеспечения, аппаратного обеспечения или их комбинации, установленных в системе, которые при работе приводят к выполнению действий системой. Одна или более компьютерных программ могут быть сконфигурированы для выполнения конкретных действий посредством включения команд, которые при исполнении устройством обработки данных приводят к выполнению действий устройством.

Хотя это техническое описание содержит много специфических подробностей реализации, их следует толковать не как ограничения в отношении объема любых изобретений или того, что может быть заявлено, а скорее как описания признаков, специфических для конкретных вариантов осуществления конкретных изобретений. Определенные признаки, описанные в данном описании в контексте отдельных вариантов осуществления, также могут быть реализованы в комбинации в одном варианте осуществления. И наоборот, различные признаки, которые описаны в контексте одного варианта осуществления, также могут быть реализованы в нескольких вариантах осуществления по отдельности или в любой подходящей подкомбинации. Кроме того, несмотря на то, что признаки могут быть описаны выше как действующие в определенных комбинациях и даже изначально заявлены как таковые, один или более признаков из заявленной комбинации в некоторых случаях могут быть исключены из комбинации, а заявленная комбинация может быть направлена на подкомбинацию или вариант подкомбинации.

Подобным образом, хотя операции изображены на графических материалах в определенном порядке, это не следует истолковывать как требование того, чтобы для достижения желаемых результатов эти операции выполнялись в конкретном показанном

порядке или в последовательном порядке или чтобы выполнялись все проиллюстрированные операции. В некоторых обстоятельствах может оказаться предпочтительной многозадачность и параллельная обработка. Кроме того, разделение различных компонентов системы в вышеописанных вариантах осуществления не следует
5 понимать как требующее такого же разделения во всех вариантах осуществления, а также следует понимать, что описанные программные компоненты и системы в целом могут быть интегрированы вместе в единый продукт программного обеспечения или упакованы в несколько продуктов программного обеспечения.

Если конкретно не указано иное, как очевидно из следующих обсуждений, то понятно,
10 что во всех обсуждениях настоящего изобретения, в которых используются такие термины, как «обработка», «вычисление», «расчет», «определение», «анализ» или т. п., они относятся к действию и/или процессам компьютера, или вычислительной системы, или подобных электронных вычислительных устройств, которые совершают манипуляции и/или преобразование данных, представленных в виде физических,
15 например электронных, величин, в другие данные, подобно представленным в виде физических величин.

Отсылка во всем настоящем описании к «одному примерному варианту осуществления», «некоторым примерным вариантам осуществления» или «примерному варианту осуществления» означает, что конкретные признак, структура или
20 характеристика, описанные в связи с примерным вариантом осуществления, включены в по меньшей мере один примерный вариант осуществления настоящего изобретения. Таким образом, появления фраз «в одном примерном варианте осуществления», «в некоторых примерных вариантах осуществления» или «в примерном варианте осуществления» в различных местах по всему настоящему описанию изобретения не
25 обязательно все относятся к одному и тому же примерному варианту осуществления. Кроме того, конкретные признаки, структуры или характеристики могут быть скомбинированы в одном или более примерных вариантах осуществления любым подходящим образом, как будет очевидно из настоящего описания изобретения специалисту в данной области техники.

30 В контексте настоящего документа, если не указано иное, использование порядковых числительных «первый», «второй», «третий» и т. д. для описания обычного объекта просто указывает на то, что производится ссылка на различные экземпляры сходных объектов, и они не предназначены для обозначения того, что объекты, описанные таким образом, должны находиться в заданной последовательности будь то во времени,
35 в пространстве, по рангу или в любом ином смысле.

В приведенной ниже формуле изобретения и в настоящем описании любой из терминов «содержащий», «состоящий из» или «который содержит» является открытым термином, что означает включение по меньшей мере следующих за ним элементов/признаков, но не исключение остальных. Поэтому термин «содержащий» при его использовании в
40 формуле изобретения не следует интерпретировать как ограничивающий в отношении средств, или элементов, или этапов, перечисляемых после него. Например, объем выражения «устройство, содержащее А и В» не следует ограничивать устройствами, состоящими только из элементов А и В. Любой из используемых в настоящем документе терминов «включающий» или «который включает» также представляет собой открытый
45 термин, который также означает включение по меньшей мере элементов/признаков, следующих за этим термином, но не исключение остальных. Таким образом, «включающий» означает «содержащий» и является его синонимом.

Следует понимать, что в приведенном выше описании примерных вариантов

осуществления настоящего изобретения различные признаки настоящего изобретения иногда группируются вместе в одном примерном варианте осуществления, на фигуре или в их описании с целью упрощения описания и содействия пониманию одного или более различных аспектов изобретения. Такой способ раскрытия изобретения, однако, 5 нельзя интерпретировать как отражающий намерение того, что формула изобретения требует большего количества признаков, чем те, которые явно перечислены в каждом пункте формулы изобретения. Вместо этого, как отражает нижеследующая формула изобретения, аспекты настоящего изобретения заключаются менее чем во всех признаках одного вышеописанного примерного варианта осуществления. Таким образом, формула 10 изобретения, следующая за разделом «Описание», настоящим явно включена в этот раздел «Описание», причем каждый пункт формулы изобретения самостоятельно представляет собой отдельный примерный вариант осуществления настоящего изобретения.

Кроме того, хотя некоторые примерные варианты осуществления, описанные в 15 настоящем документе, включают одни, а не другие признаки, включенные в другие примерные варианты осуществления, комбинации признаков из разных примерных вариантов осуществления подразумеваются как находящиеся в рамках объема настоящего изобретения и образующие разные примерные варианты осуществления, как будет понятно специалистам в данной области техники. Например, в нижеследующей 20 формуле изобретения любые заявленные примерные варианты осуществления могут применяться в любой комбинации.

В приведенном в настоящем документе описании изложено множество конкретных деталей. Однако понятно, что примерные варианты осуществления настоящего изобретения могут быть реализованы на практике без этих конкретных деталей. В 25 других случаях хорошо известные способы, структуры и технологии не показаны подробно, чтобы не затруднять понимание данного описания.

Таким образом, хотя здесь было описано то, что считается лучшими вариантами осуществления настоящего изобретения, специалистам в данной области будет понятно, что в них могут вноситься другие и дополнительные модификации без отступления от 30 идеи настоящего изобретения, и подразумевается, что все указанные изменения и модификации заявляются как находящиеся в рамках объема настоящего изобретения. Например, любые приведенные выше формулы являются только примерами процедур, которые могут использоваться. Функциональные возможности могут добавляться в структурные схемы или удаляться из них, и операции могут взаимно заменяться среди 35 функциональных блоков. Этапы могут добавляться в способы, описанные в рамках объема настоящего изобретения, или удаляться из них.

Различные аспекты и реализации настоящего изобретения можно также понять из следующих пронумерованных примерных вариантов осуществления (ППВО), которые не являются пунктами формулы изобретения.

40 ППВО 1. Способ обработки входного звукового сигнала, представленного в подполосах, для получения подвергнутого понижающему микшированию сигнала и связанных метаданных, причем способ включает:

для каждой подполосы, определение матрицы понижающего микширования и связанных метаданных; и;

45 повторное микширование каждой из указанных подполос согласно указанной матрице понижающего микширования для получения указанного подвергнутого понижающему микшированию сигнала.

ППВО 2. Способ согласно ППВО 1, в котором кодирование метаданных выполняют

с использованием одной или более стратегий квантования и одной или более стратегий кодирования с учетом ограничения целевого и/или максимального битрейта метаданных.

ППВО 3. Способ согласно ППВО 2, который включает не являющееся временно-дифференциальным энтропийное кодирование всех подполос.

5 ППВО 4. Способ согласно ППВО 3, который включает комбинирование частотного чередования с временно-дифференциальным кодированием квантованных параметров, соответствующих выбранным подполосам, для аудиокодека с низкой задержкой.

ППВО5. Способ согласно ППВО 4, который включает не являющееся энтропийным кодирование метаданных подполос.

10 ППВО 6. Способ согласно ППВО 5, который включает итеративное прохождение этапов с 3) по 5) для нахождения подходящей стратегии кодирования для выполнения требований к битрейту и качеству звучания и для уменьшения состояния декодера.

ППВО 7. Способ согласно ППВО 6, который включает уменьшение количества полос, передаваемых посредством комбинации метаданных в подполосах.

15 ППВО 8. Способ согласно ППВО 7, который включает фиксацию во времени метаданных одной или более подполос, так что метаданные подполосы не нужно передавать.

ППВО 9. Способ согласно ППВО 8, который включает использование нескольких уровней квантования для заданных метаданных, чтобы гарантировать достижение 20 целевых битрейтов.

ППВО 10. Способ согласно ППВО 9, который включает итеративное выполнение этапов ППВО 3-9 для нахождения подходящей схемы квантования для выполнения требований к битрейту и качеству звучания.

25 ППВО 11. Способ согласно ППВО 3 или ППВО 9, в котором отображение индексов из предыдущих кадров, квантованных в другое количество уровней, в индексы текущего кадра позволяет выполнять временно-дифференциальное кодирование между кадрами, не прибегая к передаче не являющегося временно-дифференциальным кадра каждый раз, когда необходим другой уровень квантования.

30 ППВО 12. Способ согласно любому из вышеуказанных ППВО, в котором квантование включает определение наилучшего значения для коэффициентов согласно текущим потребностям, путем изменения порядка вычисления и квантования последовательных коэффициентов метаданных.

ППВО 13. Система, содержащая:

один или более процессоров; и

35 постоянный машиночитаемый носитель, на котором хранятся команды, которые при исполнении одним или более процессорами приводят к выполнению одним или более процессорами операций согласно любому из ППВО 1-12.

40 ППВО 14. Постоянный машиночитаемый носитель, на котором хранятся команды, которые при исполнении одним или более процессорами приводят к выполнению одним или более процессорами операций согласно любому из ППВО 1-12.

(57) Формула изобретения

1. Способ покадрового кодирования метаданных для входного сигнала, где метаданные содержат множество по меньшей мере частично взаимосвязанных 45 параметров, вычисляемых из входного сигнала, при этом способ включает для каждого кадра:

итеративное выполнение, с использованием зацикленного процесса, следующих этапов:

выбор стратегии квантования из множества стратегий квантования для вычисления и квантования параметров;

вычисление и квантование параметров на основе выбранной стратегии квантования для получения квантованных параметров; и

5 кодирование квантованных параметров,

при этом каждая из множества стратегий квантования содержит соответствующий первый указатель, указывающий порядок, относящийся к вычислению и квантованию отдельных параметров; и

10 при этом стратегию квантования выбирают на основе целевого порогового значения битрейта,

при этом стратегия квантования, определенная для текущего кадра, отличается от стратегии квантования, определенной для предыдущего кадра; и при этом кодирование параметров включает временно-дифференциальное кодирование в разных стратегиях квантования.

15 2. Способ по п. 1, отличающийся тем, что стратегию квантования выбирают так, что битрейт закодированных квантованных параметров равен целевому пороговому значению битрейта.

3. Способ по п. 1 или п. 2, отличающийся тем, что каждая из множества стратегий квантования дополнительно содержит соответствующий второй указатель, указывающий 20 информацию для выполнения квантования параметров; и

при этом, в частности, информация для выполнения квантования параметров содержит соответствующие диапазоны квантования и/или уровни квантования для множества параметров.

4. Способ по любому из предыдущих пунктов, отличающийся тем, что кодирование 25 параметров включает временно- и/или частотно-дифференциальное кодирование.

5. Способ по любому из предыдущих пунктов, отличающийся тем, что первый указатель содержит информацию, указывающую, что:

все параметры вычисляют перед квантованием;

30 параметры вычисляют по отдельности и затем квантуют последовательно друг за другом, и по меньшей мере один параметр из множества параметров вычисляют на основе одного или более других квантованных параметров из множества параметров; или

35 все параметры вычисляют перед квантованием любого параметра; и при этом по меньшей мере один из параметров повторно вычисляют на основе другого квантованного параметра и повторно вычисленный параметр квантуют.

6. Способ по любому из предыдущих пунктов, отличающийся тем, что способ дополнительно включает, перед кодированием квантованных параметров:

отображение индексов квантованных параметров из предыдущего кадра в индексы текущего кадра.

40 7. Способ по любому из предыдущих пунктов, отличающийся тем, что зацикленный процесс включает следующие этапы:

квантование и кодирование параметров недифференциальным и/или частотно-дифференциальным образом с использованием энтропийного кодера согласно стратегии квантования;

45 оценивание первого битрейта параметров для закодированных параметров; и

выход из зацикленного процесса, если первый битрейт параметров меньше целевого порогового значения битрейта или равен ему.

8. Способ по п. 7, отличающийся тем, что зацикленный процесс дополнительно

включает следующие этапы:

если первый битрейт параметров больше целевого порогового значения битрейта: квантование и кодирование параметров недифференциальным образом без энтропии согласно стратегии квантования;

5 оценивание второго битрейта параметров для закодированных параметров; и выход из зацикленного процесса, если второй битрейт параметров меньше целевого порогового значения битрейта или равен ему.

9. Способ по п. 8, отличающийся тем, что зацикленный процесс дополнительно включает следующие этапы:

10 если второй битрейт параметров больше целевого порогового значения битрейта: квантование и кодирование параметров временно-дифференциальным образом с использованием энтропийного кодера согласно стратегии квантования;

оценивание третьего битрейта параметров для закодированных параметров; и выход из зацикленного процесса, если третий битрейт параметров меньше целевого
15 порогового значения битрейта или равен ему.

10. Способ по п. 9, отличающийся тем, что временно-дифференциальное квантование и кодирование выполняют на подмножестве параметров методом с чередованием частот относительно предыдущего кадра.

11. Способ по п. 10, отличающийся тем, что временно-дифференциальное квантование
20 и кодирование выполняют посредством циклического прохождения ряда схем временно-дифференциального кодирования с чередованием частот, так, что для каждого цикла разное подмножество параметров квантуют и кодируют временно-дифференциально, при этом остальные параметры квантуют и кодируют недифференциально.

12. Способ по любому из пп. 7-11, отличающийся тем, что выбранная стратегия
25 квантования является первой стратегией квантования, и при этом зацикленный процесс дополнительно включает:

если третий битрейт параметров больше целевого порогового значения битрейта: выбор второй стратегии квантования из множества стратегий квантования так, что
30 ожидается, что при применении второй стратегии квантования битрейт будет меньше битрейта использования первой стратегии квантования; и

повторение этапов зацикленного процесса по п. 7.

13. Способ по любому из пп. 7-11, отличающийся тем, что параметры представлены
в первом количестве полос частот, и при этом зацикленный процесс дополнительно
включает следующие этапы:

35 если третий битрейт параметров больше целевого порогового значения битрейта: уменьшение количества полос частот, представляющих параметры, до второго количества, которое меньше первого количества, так, что общее количество параметров, которые необходимо квантовать и кодировать, уменьшается; или

повторное использование параметров в одной или более полосах частот из
40 предыдущего кадра в текущем кадре; и

повторение этапов зацикленного процесса по п. 7.

14. Способ по п. 12 или п. 13, отличающийся тем, что зацикленный процесс
дополнительно включает следующие этапы:

перед выбором второй стратегии квантования, или уменьшением количества полос
45 частот, или повторным использованием параметров:

получение минимума среди первого, второго и третьего битрейтов параметров; и выход из зацикленного процесса, если минимум меньше максимального порогового значения битрейта или равен ему.

15. Способ по любому из предыдущих пунктов, отличающийся тем, что способ применяют к кодированию метаданных кодека голосовых и звуковых служб с эффектом присутствия, IVAS, или кодека амбиофонии.

5 16. Устройство покадрового кодирования метаданных для входного сигнала, содержащее процессор и устройство памяти, соединенное с процессором, при этом процессор приспособлен вызывать выполнение устройством способа по любому из предыдущих пунктов.

10 17. Машиночитаемый носитель данных, на котором хранится программа, содержащая команды, которые при исполнении процессором вызывают выполнение процессором способа по любому из пп. 1-15.

15

20

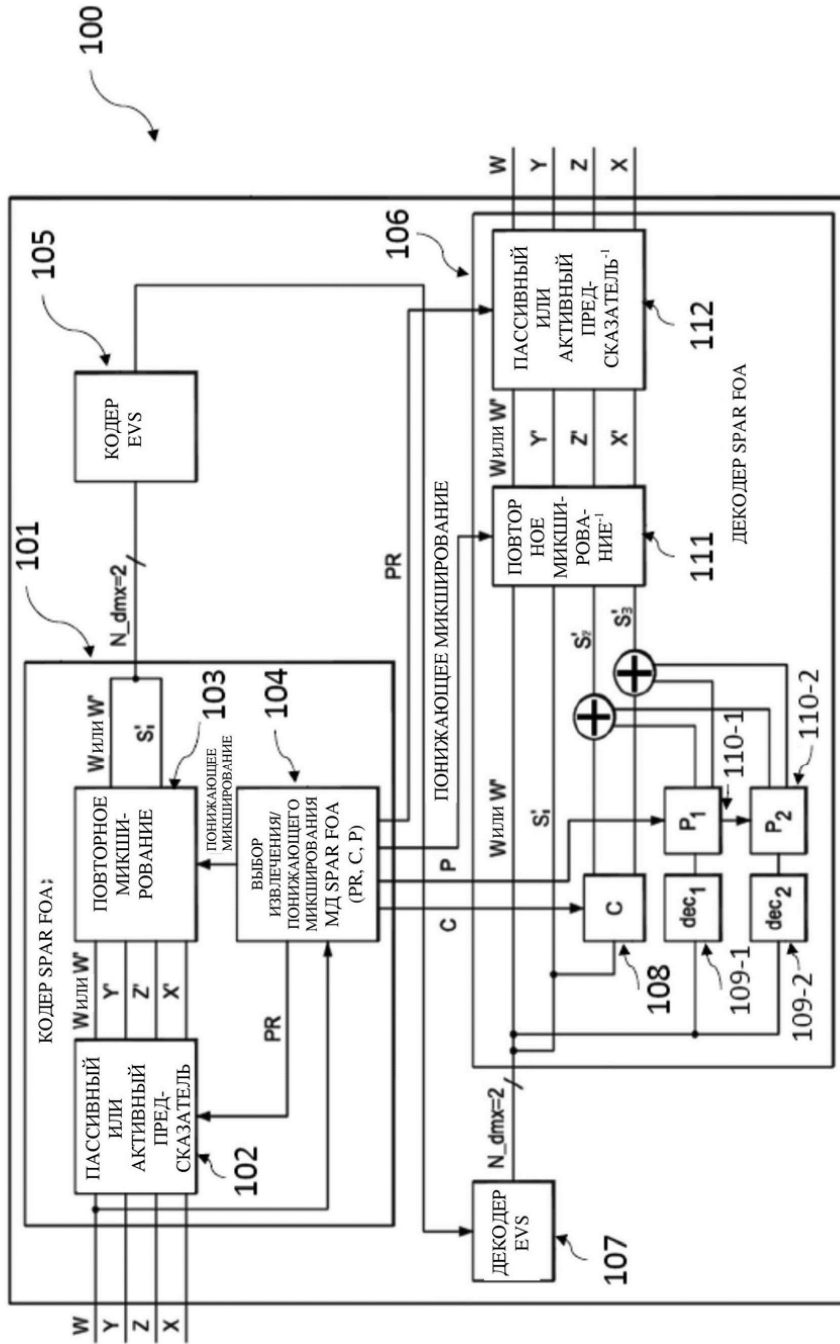
25

30

35

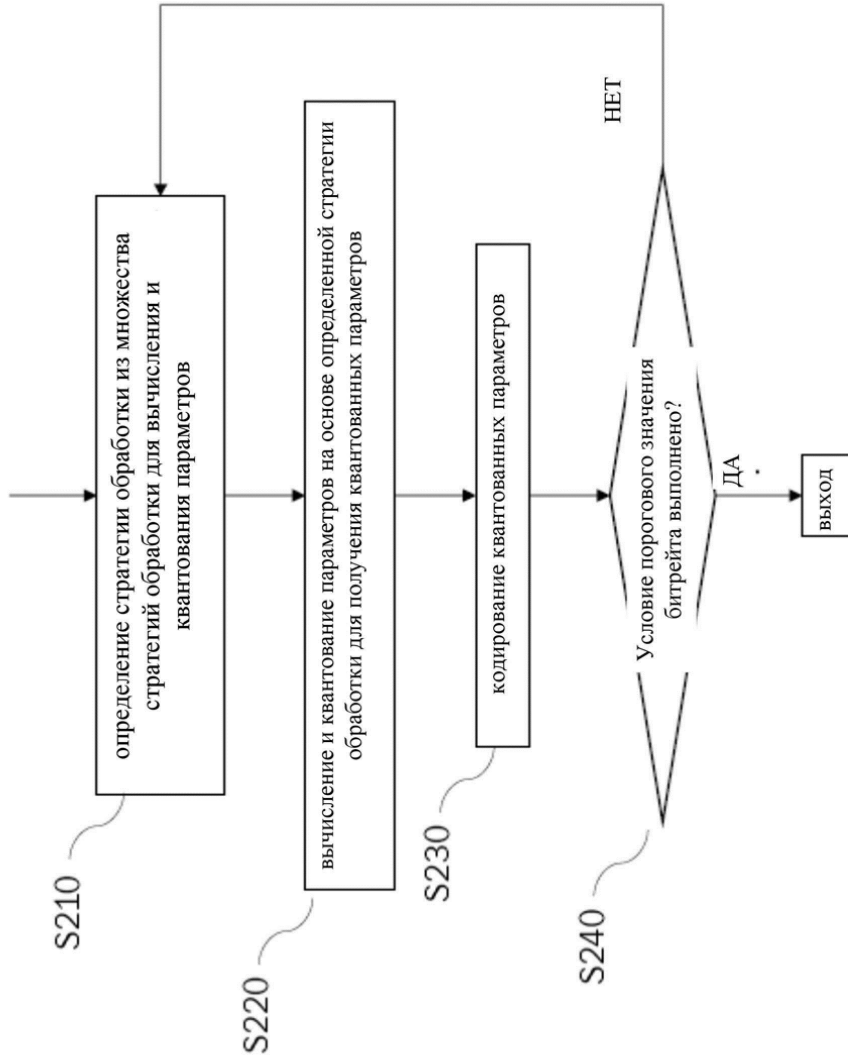
40

45

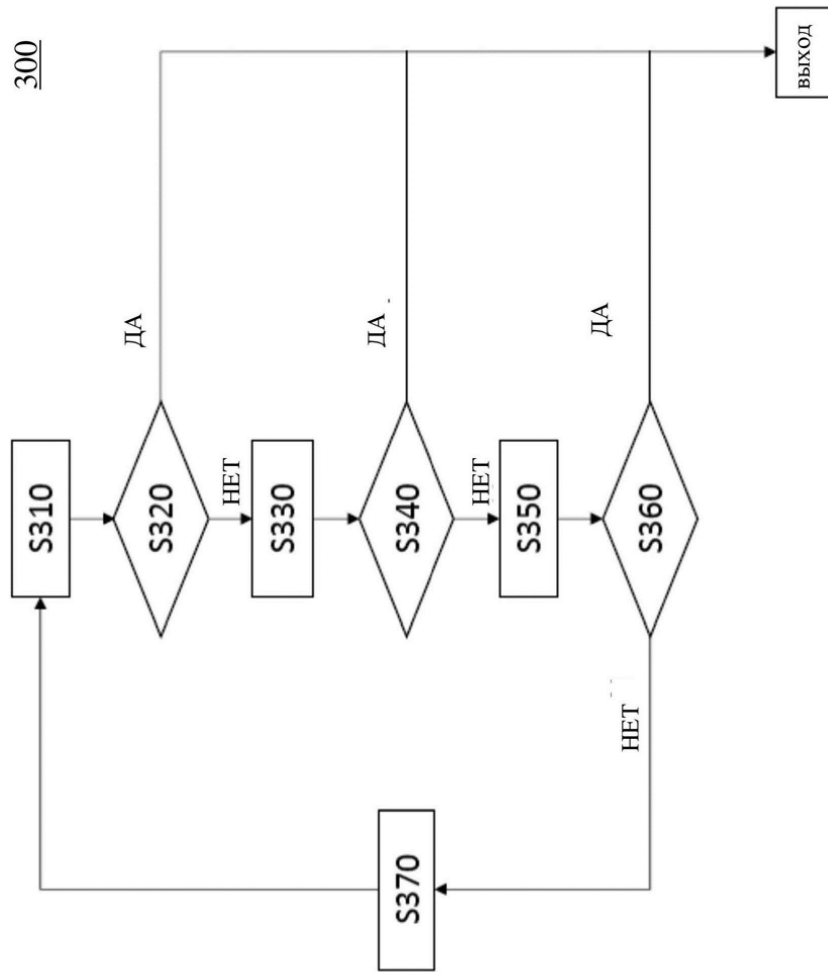


Фиг. 1

200

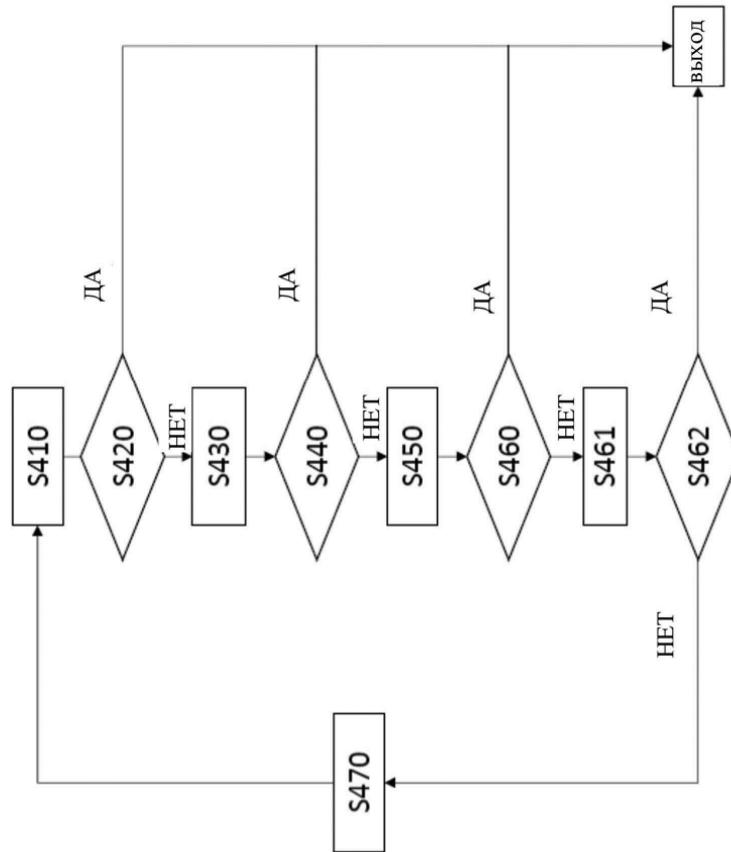


Фиг. 2



Фиг. 3

400



Фиг. 4