



ФЕДЕРАЛЬНАЯ СЛУЖБА
ПО ИНТЕЛЛЕКТУАЛЬНОЙ СОБСТВЕННОСТИ

(12) ОПИСАНИЕ ИЗОБРЕТЕНИЯ К ПАТЕНТУ

(52) СПК

G10L 19/008 (2006.01); H03M 7/30 (2006.01)

(21)(22) Заявка: 2016105613, 15.07.2014

(24) Дата начала отсчета срока действия патента:
15.07.2014

Дата регистрации:
02.03.2018

Приоритет(ы):

(30) Конвенционный приоритет:

22.07.2013 EP EP13177353;

22.07.2013 EP EP13177350;

22.07.2013 EP EP13177348;

22.07.2013 EP EP13177346;

18.10.2013 EP EP13189366

(43) Дата публикации заявки: 28.08.2017 Бюл. № 25

(45) Опубликовано: 02.03.2018 Бюл. № 7

(85) Дата начала рассмотрения заявки РСТ на
национальной фазе: 24.02.2016

(86) Заявка РСТ:
EP 2014/065106 (15.07.2014)

(87) Публикация заявки РСТ:
WO 2015/010947 (29.01.2015)

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

(72) Автор(ы):

ДИШ Саша (DE),

НАГЕЛЬ Фредерик (DE),

ГАЙГЕР Ральф (DE),

ТХОШКАХНА Баладжи Нагендран (DE),

ШМИДТ Константин (DE),

БАЙЕР Штефан (DE),

НОЙКАМ Кристиан (DE),

ЭДЛЕР Бернд (DE),

ХЕЛЬМРИХ Кристиан (DE)

(73) Патентообладатель(и):

ФРАУНХОФЕР-ГЕЗЕЛЛЬШАФТ ЦУР

ФЕРДЕРУНГ ДЕР АНГЕВАНДТЕН

ФОРШУНГ Е.Ф. (DE)

(56) Список документов, цитированных в отчете
о поиске: US 7328161 B2, 05.02.2008. US
7460990 B2, 02.012.2008. RU 2325708 C2,
27.05.2008. RU 2323469 C2, 27.04.2008. RU
2422922 C1, 27.06.2011. US 2006/0031075 A1,
09.02.2006. US 7756713 B2, 13.07.2010. US
8473301 B2, 25.06.2013.

(54) АУДИОКОДЕР, АУДИОДЕКОДЕР И СВЯЗАННЫЕ СПОСОБЫ С ИСПОЛЬЗОВАНИЕМ
ДВУХКАНАЛЬНОЙ ОБРАБОТКИ В ИНФРАСТРУКТУРЕ ИНТЕЛЛЕКТУАЛЬНОГО ЗАПОЛНЕНИЯ
ИНТЕРВАЛОВ ОТСУТСТВИЯ СИГНАЛА

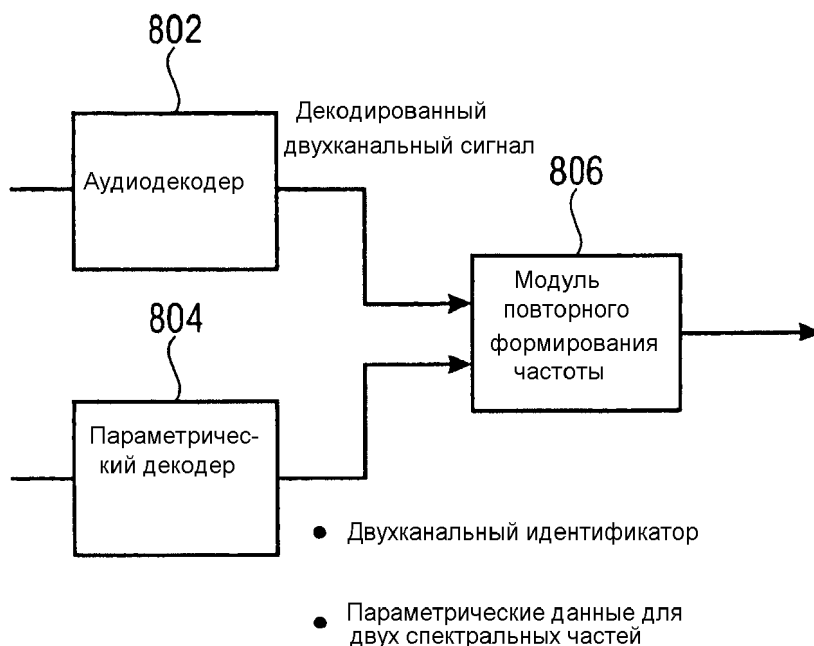
(57) Реферат:

Изобретение относится к кодированию/
декодированию аудио сигналов, в частности к
кодированию аудиосигналов с использованием
интеллектуального заполнения интервалов (IGF).
Технический результат – обеспечение
расширенной полосы пропускания аудиосигнала
на низких скоростях передачи битов за счет

повышения эффективности сжатия. Устройство
для формирования декодированного
двухканального сигнала содержит:
аудиопроцессор для декодирования
кодированного двухканального сигнала для того,
чтобы получать первый набор первых
спектральных частей; параметрический декодер

для предоставления параметрических данных для второго набора вторых спектральных частей и двухканального идентификатора, идентифицирующего первое или второе другое двухканальное представление для вторых спектральных частей; и модуль повторного формирования частоты для повторного

формирования второй спектральной части в зависимости от первой спектральной части первого набора первых спектральных частей, параметрических данных для второй части и двухканального идентификатора для второй части. 6 н. и 16 з.п. ф-лы, 41 ил.



Фиг. 8а



FEDERAL SERVICE
FOR INTELLECTUAL PROPERTY

(12) **ABSTRACT OF INVENTION**

(52) CPC

G10L 19/008 (2006.01); *H03M 7/30* (2006.01)(21)(22) Application: **2016105613, 15.07.2014**(24) Effective date for property rights:
15.07.2014Registration date:
02.03.2018

Priority:

(30) Convention priority:
22.07.2013 EP EP13177353;
22.07.2013 EP EP13177350;
22.07.2013 EP EP13177348;
22.07.2013 EP EP13177346;
18.10.2013 EP EP13189366(43) Application published: **28.08.2017 Bull. № 25**(45) Date of publication: **02.03.2018 Bull. № 7**(85) Commencement of national phase: **24.02.2016**(86) PCT application:
EP 2014/065106 (15.07.2014)(87) PCT publication:
WO 2015/010947 (29.01.2015)

Mail address:

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

(72) Inventor(s):

DISH Sasha (DE),
NAGEL Frederik (DE),
GAJGER Ralf (DE),
TKHOSHKAKHNA Baladzhi Nagendran (DE),
SHMIDT Konstantin (DE),
BAJER Shtefan (DE),
NOJKAM Kristian (DE),
EDLER Bernd (DE),
KHELMRIKH Kristian (DE)

(73) Proprietor(s):

FRAUNKHOFER-GEZELLSCHAFT TSUR
FERDERUNG DER ANGEVANDTEN
FORSHUNG E.F. (DE)(54) **AUDIO CODER, AUDIO DECODER AND RELATED METHODS USING TWO-CHANNEL PROCESSING IN INFRASTRUCTURE OF INTELLECTUAL FILLING INTERVALS OF SIGNAL ABSENCE**

(57) Abstract:

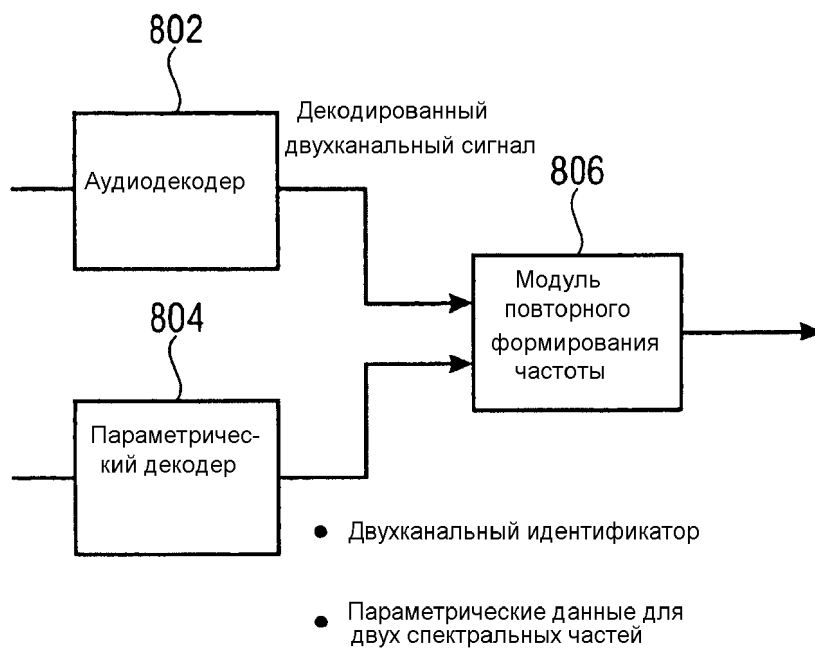
FIELD: physics.

SUBSTANCE: device for generating a decoded two-channel signal comprises: an audio processor for decoding the encoded two-channel signal in order to obtain the first set of the first spectral portions; a parametric decoder for providing parametric data for the second set of the second spectral portions and a two-channel identifier identifying the first or the second other two-channel representation for the second spectral

portions; and a frequency re-shaping unit for re-forming the second spectral portion, depending on the first spectral portion of the first set of the first spectral portions, the parametric data for the second portion, and the two-channel identifier for the second portion.

EFFECT: providing an extended bandwidth of the audio signal at low bit rates due to the increased compression efficiency.

22 cl, 41 dwg



Фиг. 8а

Подробное описание изобретения

Настоящее изобретение относится к кодированию/декодированию аудио и, в частности, к кодированию аудио с использованием интеллектуального заполнения интервалов (IGF).

Кодирование аудио представляет собой область сжатия сигналов, которая связана с использованием избыточности и нерелевантности в аудиосигналах с использованием психоакустических сведений. На сегодняшний день аудиокодекам типично требуется приблизительно 60 Кбит/с/канал для перцепционно прозрачного кодирования практически любого типа аудиосигнала. Более новые кодеки нацелены на уменьшение скорости передачи битов при кодировании посредством использования спектральных подобий в сигнале с использованием таких технологий, как расширение полосы пропускания (BWE). BWE-схема использует набор параметров для низкой скорости передачи битов, чтобы представлять высокочастотные (HF) компоненты аудиосигнала. HF-спектр заполнен спектральным содержимым из низкочастотных (LF) областей, и спектральная форма, наклон и временная непрерывность регулируются для того, чтобы поддерживать тембр и цвет исходного сигнала. Такие BWE-способы позволяют аудиокодекам сохранять хорошее качество даже на низких скоростях передачи битов приблизительно в 24 Кбит/с/канал.

Хранение или передача аудиосигналов зачастую подчиняются строгим ограничениям скорости передачи битов. В прошлом, кодеры принудительно существенно уменьшали полосу пропускания передаваемого аудиосигнала, когда была доступна только очень низкая скорость передачи битов.

Современные аудиокодеки в наше время могут кодировать широкополосные сигналы посредством использования способов расширения полосы пропускания (BWE)[1]. Эти алгоритмы основываются на параметрическом представлении высокочастотного содержимого (HF), который формируется из кодированной на основе формы сигналов низкочастотной части (LF) декодированного сигнала, посредством транспозиции в спектральную HF-область ("наложения") и применения постобработки на основе параметров. В BWE-схемах, восстановление спектральной HF-области выше данной так называемой частоты разделения зачастую основано на спектральном наложении. Типично, HF-область состоит из нескольких смежных наложений, и каждое из этих наложений получается из полосовых (BP) областей LF-спектра ниже данной частоты разделения. Системы предшествующего уровня техники эффективно выполняют наложение в представлении на основе гребенки фильтров, например, гребенки квадратурных зеркальных фильтров (QMF), посредством копирования набора смежных подполосных коэффициентов из исходной в целевую область.

Еще одна технология, разработанная в современных аудиокодеках, которая повышает эффективность сжатия и за счет этого обеспечивает расширенную полосу пропускания аудиосигнала на низких скоростях передачи битов, представляет собой синтетическую замену на основе параметров подходящих частей спектров звука. Например, шумоподобные части сигнала исходного аудиосигнала могут быть заменены без существенных потерь субъективного качества посредством искусственного шума, сформированного в декодере, и масштабированы посредством параметров вспомогательной информации. Один пример представляет собой инструментальное средство для перцепционного замещения шума (PNS), содержащееся в усовершенствованном кодировании аудио (AAC) на основе MPEG-4 [5].

Дополнительная мера, которая также обеспечивает расширенную полосу пропускания аудиосигнала на низких скоростях передачи битов, представляет собой технологию

заполнения шумом, содержащуюся в стандартизированном кодировании речи и аудио (USAC) на основе MPEG-D [7]. Интервалы отсутствия сигнала в спектре (нули), которые логически выводятся посредством мертвой зоны квантователя вследствие слишком приблизительного квантования, затем заполняются искусственным шумом в декодере и масштабируются посредством постобработки на основе параметров.

Другая система предшествующего уровня техники называется "точной спектральной заменой (ASR)" [2-4]. В дополнение к кодеку на основе формы сигналов, ASR использует выделенную стадию синтеза сигналов, которая восстанавливает перцепционно важные синусоидальные части сигнала в декодере. Кроме того, система, описанная в [5], основывается на синусоидальном моделировании в HF-области кодера на основе формы сигналов, чтобы обеспечивать расширенную полосу пропускания аудиосигнала, имеющую неплохое перцепционное качество на низких скоростях передачи битов. Все эти способы включают в себя преобразование данных во второй области, отличное от модифицированного дискретного косинусного преобразования (MDCT), а также довольно комплексные стадии анализа/синтеза для сохранения синусоидальных HF-компонентов.

Фиг. 13а иллюстрирует принципиальную схему аудиокодера для технологии расширения полосы пропускания, например, используемой при высокоэффективном усовершенствованном кодировании аудио (HE-AAC). Аудиосигнал в линии 1300 вводится в систему фильтров, состоящую из нижних частот 1302 и верхних частот 1304. Сигнал, выводимый посредством фильтра 1304 верхних частот, вводится в модуль 1306 извлечения/кодирования параметров. Модуль 1306 извлечения/кодирования параметров выполнен с возможностью вычисления и кодирования параметров, таких как, например, параметр спектральной огибающей, параметр добавления шума, параметр пропущенных гармоник или параметр обратной фильтрации. Эти извлеченные параметры вводятся в мультиплексор 1308 потоков битов. Выходной сигнал нижних частот вводится в процессор, типично содержащий функциональность модуля 1310 понижающей дискретизации и базового кодера 1312. Нижние частоты 1302 ограничивают полосу пропускания, которая должна кодироваться, значительно меньшей полосой пропускания, чем возникающая исходном входном аудиосигнале на линии 1300. Это предоставляет значительное усиление при кодировании вследствие того факта, что полные функциональности, осуществляемые в базовом декодере, должны работать только для сигнала с уменьшенной полосой пропускания. Когда, например, полоса пропускания аудиосигнала на линии 1300 составляет 20 кГц, и когда фильтр 1302 нижних частот примерно имеет полосу пропускания в 4 кГц, чтобы удовлетворять теореме дискретизации, теоретически достаточно того, что сигнал после модуля понижающей дискретизации имеет частоту дискретизации в 8 кГц, что является существенным уменьшением по сравнению с частотой дискретизации, требуемой для аудиосигнала 1300, которая должна составлять, по меньшей мере, 40 кГц.

Фиг. 13b иллюстрирует принципиальную схему соответствующего декодера расширения полосы пропускания. Декодер содержит мультиплексор 1320 потоков битов. Демультимплексор 1320 потоков битов извлекает входной сигнал для базового декодера 1322 и входной сигнал для декодера 1324 параметров. Выходной сигнал базового декодера имеет, в вышеприведенном примере, частоту дискретизации в 8 кГц, и следовательно, полосу пропускания в 4 кГц, тогда как для восстановления полной полосы пропускания выходной сигнал модуля 1330 восстановления высоких частот должен иметь 20 кГц, что требует частоты дискретизации, по меньшей мере, в 40 кГц. Для обеспечения возможности этого, требуется процессор декодера, имеющий

функциональность модуля 1325 повышающей дискретизации и гребенки 1326 фильтров. Модуль 1330 восстановления высоких частот затем принимает частотно проанализированный низкочастотный сигнал, выводимый посредством гребенки 1326 фильтров, и восстанавливает частотный диапазон, заданный посредством фильтра 1304 верхних частот по фиг. 13а, с использованием параметрического представления полосы высоких частот. Модуль 1330 восстановления высоких частот имеет несколько функциональностей, таких как повторное формирование диапазона верхних частот с использованием исходного диапазона в диапазоне низких частот, регулирование спектральной огибающей, функциональность добавления шума и функциональность для того, чтобы вводить пропущенные гармоники в диапазоне верхних частот, и если применяется и вычисляется в кодере по фиг. 13а, операция обратной фильтрации, чтобы учитывать тот факт, что диапазон верхних частот типично не является настолько тональным, как диапазон нижних частот. В HE-AAC, пропущенные гармоники повторно синтезируются на стороне декодера и размещаются точно в середине полосы частот восстановления. Следовательно, все линии пропущенных гармоник, которые определяются в определенной полосе частот восстановления, не размещены в значениях частоты, в которых они располагаются в исходном сигнале. Вместо этого, эти линии пропущенных гармоник размещены в частотах в центре определенной полосы частот. Таким образом, когда линия пропущенных гармоник в исходном сигнале размещена очень близко к границе полосы частот восстановления в исходном сигнале, ошибка в частоте, введенная посредством размещения этой линии пропущенных гармоник в восстановленном сигнале в центре полосы частот, находится близко к 50% отдельной полосы частот восстановления, для которой сформированы и переданы параметры.

Кроме того, даже если типичные аудио базовые кодеры работают в спектральной области, базовый декодер, тем не менее, формирует сигнал временной области, который затем снова преобразуется в спектральную область посредством функциональности гребенки 1326 фильтров. Это вводит дополнительные задержки при обработке, может вводить артефакты вследствие тандемной обработки преобразования сначала из спектральной области в частотную область и снова преобразования типично в другую частотную область, и, конечно, это также требует значительной сложности вычислений и в силу этого электроэнергии, что представляет собой проблему, в частности, когда технология расширения полосы пропускания применяется в мобильных устройствах, к примеру, в мобильных телефонах, планшетных или переносных компьютерах и т.д.

Современные аудиокодеки выполняют кодирование аудио с низкой скоростью передачи битов с использованием BWE в качестве неотъемлемой части схемы кодирования. Тем не менее, BWE-технологии ограничены тем, что они заменяют только высокочастотный (HF) спектр. Более того, они не обеспечивают возможность кодирования на основе формы сигналов перцепционно важного содержимого выше данной частоты разделения. Следовательно, современные аудиокодеки теряют HF-детали или тембр, когда реализуется BWE, поскольку точное совмещение тональных гармоник сигнала не учитывается в большинстве систем.

Другой недостаток BWE-систем современного уровня техники заключается в необходимости преобразования аудиосигнала в новую область для реализации BWE (например, преобразования из MDCT-в QMF-область). Это приводит к усложнению синхронизации, дополнительной вычислительной сложности и повышенным требованиям к запоминающему устройству.

В случае двухканальных пар, по существу предусмотрено несколько канальных представлений, таких как объединенное канальное представление или отдельное

канальное представление. Известное объединенное представление является средним/боковым представлением, в котором средний канал является суммой левого и правого канала, и в котором боковой канал является разностью между левым и правым каналом.

Другое представление представляет собой канал низведения и остаточный канал и
5 дополнительный коэффициент прогнозирования, который дает возможность воссоздавать левый и правый канал из низведения и остатка. Отдельное представление, в этом случае, должно представлять собой левый и правый отдельный канал или, в общем, первый канал и второй канал.

Кроме того, возникает ситуация, в которой исходный диапазон для операций
10 заполнения интервалов отсутствия сигнала может демонстрировать сильную корреляцию, в то время как целевой диапазон не демонстрирует эту сильную корреляцию. Когда исходный диапазон, в этом варианте осуществления, кодируется с использованием первого стереопредставления, такого как среднее/боковое представление, чтобы уменьшать скорость передачи битов для базовой частотной
15 части, в таком случае неправильное двухканальное изображение формируется для части восстановления или целевого диапазона. С другой стороны, когда исходный диапазон не демонстрирует корреляцию либо имеет только небольшую корреляцию, и целевой диапазон имеет небольшую корреляцию либо не имеет корреляции, с другой стороны, операция прямого заполнения интервалов отсутствия сигнала приводит к артефактам.

20 Следовательно, цель настоящего изобретения заключается в том, чтобы предоставлять усовершенствованный принцип кодирования/декодирования для двухканальных представлений.

Это цель достигается посредством аудиодекодера по п. 1, аудиокодера по п. 14, способа декодирования аудио по п. 19, способа декодирования аудио по п. 20 или
25 компьютерной программы по п. 21.

Настоящее изобретение основано на таких выявленных сведениях, что ситуация корреляции является важной не только для исходного диапазона, но также и для целевого диапазона. Кроме того, настоящее изобретение подтверждает такую ситуацию, что различные ситуации корреляции могут возникать в исходном диапазоне и целевом
30 диапазоне. Когда, например, рассматривается речевой сигнал с высокочастотным шумом, может возникать такая ситуация, что полоса низких частот, содержащая речевой сигнал с небольшим числом обертонов, имеет высокую корреляцию в левом канале и правом канале, когда динамик размещен посередине. Тем не менее, часть высоких частот может иметь сильную декорреляцию вследствие того факта, что может возникать
35 отличный высокочастотный шум с левой стороны по сравнению с другим высокочастотным шумом или отсутствием высокочастотного шума с правой стороны. Таким образом, когда должна выполняться операция прямого заполнения интервалов отсутствия сигнала, которая игнорирует эту ситуацию, в таком случае часть высоких частот также должна быть коррелирована, и это может формировать серьезные
40 артефакты пространственной сегрегации в восстановленном сигнале. Чтобы разрешать эту проблему, вычисляются параметрические данные для полосы частот восстановления или, в общем, для второго набора вторых спектральных частей, которые должны быть восстановлены с использованием первого набора первых спектральных частей для того, чтобы идентифицировать первое или второе другое двухканальное представление
45 для второй спектральной части или, другими словами, для полосы частот восстановления. Следовательно, на стороне кодера двухканальный идентификатор вычисляется для вторых спектральных частей, т.е. для частей, для которых, дополнительно, вычисляется информация энергии для полос частот восстановления.

Модуль повторного формирования частоты на стороне декодера затем повторно формирует вторую спектральную часть в зависимости от первой части из первого набора первых спектральных частей, т.е. исходного диапазона, и параметрических данных для второй части, таких как информация энергии спектральной огибающей или любые другие данные спектральной огибающей, и дополнительно, в зависимости от двухканального идентификатора для второй части, т.е. для этой повторно рассматриваемой полосы частот восстановления.

Двухканальный идентификатор предпочтительно передается в качестве флага для каждой полосы частот восстановления, и эти данные передаются из кодера в декодер, и декодер затем декодирует базовый сигнал, как указано посредством предпочтительно вычисленных флагов для полос базовых частот. Затем в реализации, базовый сигнал сохраняется в обоих стереопредставлениях (например, левый/правый и средний/боковой), и для заполнения частотными IGF-фрагментами представление исходных фрагментов выбрано таким образом, что оно соответствует представлению целевых фрагментов, как указано посредством флагов двухканального идентификатора для интеллектуального заполнения интервалов отсутствия сигнала или полос частот восстановления, т.е. для целевого диапазона.

Следует подчеркнуть, что эта процедура работает не только для стереосигналов, т.е. для левого канала и правого канала, но также и работает для многоканальных сигналов. В случае многоканальных сигналов, таким способом могут обрабатываться несколько пар различных каналов, к примеру, левый и правый канал в качестве первой пары, левый канал объемного звучания и правый объемного звучания в качестве второй пары и центральный канал и LFE-канала в качестве третьей пары. Другие спаривания могут определяться для форматов с более высоким числом выходных каналов, к примеру, 7.1, 11.1 и т.д.

Дополнительный аспект основан на таких выявленных сведениях, что проблемы, связанные с разделением расширения полосы пропускания, с одной стороны, и базового кодирования, с другой стороны, могут разрешаться и преодолеваются посредством выполнения расширения полосы пропускания в той спектральной области, в которой работает базовый декодер. Следовательно, предоставляется полноскоростной базовый декодер, который кодирует и декодирует полный диапазон аудиосигнала. Это не требует модуля понижающей дискретизации на стороне кодера и модуля повышающей дискретизации на стороне декодера. Вместо этого, вся обработка выполняется в области полной частоты дискретизации или полной полосы пропускания. Чтобы получить высокое усиление при кодировании, аудиосигнал анализируется для того, чтобы находить первый набор первых спектральных частей, который должен быть кодирован с высоким разрешением, причем этот первый набор первых спектральных частей может включать в себя, в варианте осуществления, тональные части аудиосигнала. С другой стороны, нетональные или зашумленные компоненты в аудиосигнале, составляющем второй набор вторых спектральных частей, параметрически кодируются с низким спектральным разрешением. Кодированный аудиосигнал в таком случае требует только первого набора первых спектральных частей, кодированных с сохранением формы сигнала с помощью высокого спектрального разрешения, и дополнительно, второго набора вторых спектральных частей, кодированных параметрически с низким разрешением с использованием частотных "фрагментов", получаемых из первого набора. На стороне декодера, базовый декодер, который представляет собой полнополосный декодер, восстанавливает первый набор первых спектральных частей с сохранением формы сигнала, т.е. без сведений о том, что приспособливается

дополнительное повторное формирование частоты. Тем не менее, за счет этого сформированный спектр имеет множество интервалов в спектре. Эти интервалы отсутствия сигнала затем заполнены с помощью изобретаемой технологии интеллектуального заполнения интервалов (IGF) посредством использования повторного формирования частоты, применяющего параметрические данные, с одной стороны, и использования исходного спектрального диапазона, т.е. первых спектральных частей, восстановленных посредством полноскоростного аудиодекодера, с другой стороны.

В дополнительных вариантах осуществления, спектральные части, которые восстановлены только посредством заполнения шумом, а не репликации полосы пропускания или заполнения частотными фрагментами, составляют третий набор третьих спектральных частей. Вследствие того факта, что принцип кодирования работает в одной области для базового кодирования/декодирования, с одной стороны, и повторного формирования частоты, с другой стороны, IGF ограничен не только заполнением диапазона верхних частот, но может заполнять диапазоны нижних частот, либо посредством заполнения шумом без повторного формирования частоты, либо посредством повторного формирования частоты с использованием частотного фрагмента в другом частотном диапазоне.

Кроме того, следует подчеркнуть, что информация относительно спектральных энергий, информация относительно отдельных энергий (или информация отдельных энергий), информация относительно энергии выживания (или информация энергии выживания), информация относительно энергии фрагментов (или информация энергии фрагментов) либо информация относительно недостающей энергии (или информация недостающей энергии) может содержать не только значение энергии, но также и (например, абсолютное) значение амплитуды, значение уровня или любое другое значение, из которого может быть получено конечное значение энергии. Следовательно, информация относительно энергии, например, может содержать само значение энергии и/или значение уровня и/или амплитуды, и/или абсолютной амплитуды.

Дополнительный аспект основан на таких выявленных сведениях, что качество звука восстановленного сигнала может повышаться через IGF, поскольку весь спектр является доступным для базового кодера, так что, например, перцепционно важные тональные части в высоком спектральном диапазоне по-прежнему могут кодироваться посредством базового кодера, а не параметрической замены. Дополнительно, выполняется операция заполнения интервалов отсутствия сигнала с использованием частотных фрагментов из первого набора первых спектральных частей, который представляет собой, например, набор тональных частей типично из диапазона нижних частот, но также и из диапазона верхних частот при наличии. Тем не менее, для регулирования спектральной огибающей на стороне декодера, спектральные части из первого набора спектральных частей, расположенных в полосе частот восстановления, дополнительно не постобрабатываются, например, посредством регулирования спектральной огибающей. Только оставшиеся спектральные значения в полосе частот восстановления, которые не исходят из базового декодера, должны быть подвергнуты регулированию огибающей с использованием информации огибающей. Предпочтительно, информация огибающей является информацией полнополосной огибающей с учетом энергии первого набора первых спектральных частей в полосе частот восстановления и второго набора вторых спектральных частей в идентичной полосе частот восстановления, причем вторые спектральные значения во втором наборе вторых спектральных частей указываются как равные нулю и в силу этого не кодируются посредством базового кодера, а параметрически кодируются с помощью информации энергии низкого разрешения.

Обнаружено, что абсолютные значения энергии, нормализованные относительно полосы пропускания соответствующей полосы частот или не нормализованные, являются полезными и очень эффективными при применении на стороне декодера. Это, в частности, применимо, когда коэффициенты усиления должны вычисляться на основе

5 остаточной энергии в полосе частот восстановления, недостающей энергии в полосе частот восстановления и информации частотных фрагментов в полосе частот восстановления.

Кроме того, предпочтительно, чтобы кодированный поток битов покрывал не только информацию энергии для полос частот восстановления, но и, дополнительно,

10 коэффициенты масштабирования для полос частот коэффициентов масштабирования, расширяющихся вплоть до максимальной частоты. Это обеспечивает то, что для каждой полосы частот восстановления, для которой доступна определенная тональная часть, т.е. первая спектральная часть, этот первый набор первой спектральной части может фактически декодироваться с правильной амплитудой. Кроме того, в дополнение к

15 коэффициенту масштабирования для каждой полосы частот восстановления, энергия для этой полосы частот восстановления формируется в кодере и передается в декодер. Кроме того, предпочтительно, чтобы полосы частот восстановления совпадали с полосами частот коэффициентов масштабирования, либо в случае энергетической группировки, по меньшей мере, границы полосы частот восстановления совпадали с

20 границами полос частот коэффициентов масштабирования.

Дополнительный аспект основан на таких выявленных сведениях, что определенные ухудшения качества звука могут быть исправлены посредством применения схемы сигнально-адаптивного заполнения частотными фрагментами. С этой целью, анализ на стороне кодера выполняется для того, чтобы выявлять возможный вариант наилучше

25 совпадающей исходной области для определенной целевой области. Информация совпадения, идентифицирующая для целевой области определенную исходную область, вместе с необязательно некоторой дополнительной информацией, формируется и передается в качестве вспомогательной информации в декодер. Декодер затем применяет операцию заполнения частотными фрагментами с использованием информации

30 совпадения. С этой целью, декодер считывает информацию совпадения из передаваемого потока данных или файла данных и осуществляет доступ к исходной области, идентифицированной для определенной полосы частот восстановления и, если указывается в информации совпадения, дополнительно выполняет некоторую обработку этих данных исходной области, чтобы формировать необработанные спектральные

35 данные для полосы частот восстановления. Затем этому результату операции заполнения частотными фрагментами, т.е. необработанным спектральным данным для полосы частот восстановления, придается определенная форма с использованием информации спектральной огибающей, чтобы, в завершение, получать полосу частот восстановления, которая также содержит первые спектральные части, к примеру, тональные части. Тем

40 не менее, эти тональные части не формируются посредством схемы адаптивного заполнения фрагментами, а эти первые спектральные части выводятся посредством самого аудиодекодера или базового декодера.

Схема адаптивного выбора спектральных фрагментов может работать с низкой степенью детализации. В этой реализации, исходная область подразделяется на типично

45 перекрывающиеся исходные области и целевую область, или полосы частот восстановления задаются посредством неперекрывающихся целевых частотных областей. Далее подобия между каждой исходной областью и каждой целевой областью определяются на стороне кодера, и наилучше совпадающая пара исходной области и

целевой области идентифицирована посредством информации совпадения, и на стороне декодера, исходная область, идентифицированная в информации совпадения, используется для формирования необработанных спектральных данных для полосы частот восстановления.

5 В целях получения большей степени детализации каждой исходной области разрешается сдвигаться, чтобы получать определенное запаздывание, при котором подобия являются максимальными. Это запаздывание может быть идентичным по точности элементу разрешения по частоте и обеспечивает возможность еще лучшего совпадения между исходной областью и целевой областью.

10 Кроме того, в дополнение к только идентификации наилучше совпадающей пары, это запаздывание корреляции также может передаваться в информации совпадения и, дополнительно, даже знак может передаваться. Когда знак определяется как отрицательный на стороне кодера, то соответствующий флаг знака также передается в информации совпадения, и на стороне декодера, спектральные значения исходной
15 области умножаются на -1 либо, в комплексном представлении, "циклически сдвигаются" на 180 градусов.

Дополнительная реализация этого изобретения применяет операцию отбеливания фрагментов. Отбеливание спектра удаляет приблизительную информацию спектральной огибающей и подчеркивает точную спектральную структуру, которая представляет
20 главный интерес для оценки подобия фрагментов. Следовательно, частотный фрагмент, с одной стороны, и/или исходный сигнал, с другой стороны, отбеливаются до вычисления меры взаимной корреляции. Когда только фрагмент отбелен с использованием предварительно заданной процедуры, передается флаг отбеливания, указывающий декодеру то, что идентичный предварительно заданный процесс отбеливания должен
25 применяться к частотному фрагменту в IGF.

Относительно выбора фрагментов, предпочтительно использовать запаздывание корреляции, чтобы спектрально сдвигать повторно сформированный спектр посредством целого числа элементов выборки преобразования. В зависимости от базового преобразования, спектральный сдвиг может требовать коррекций с
30 суммированием. В случае нечетных запаздываний фрагмент дополнительно модулирован через умножение на переменную временную последовательность из $-1/1$ для того, чтобы компенсировать представление с обратной частотой каждой второй полосы частот в MDCT. Кроме того, знак результата корреляции применяется при формировании частотного фрагмента.

35 Кроме того, предпочтительно использовать отсечение и стабилизацию фрагментов, чтобы удостовериться, что не допускаются артефакты, созданные посредством быстро изменяющихся исходных областей для идентичной области восстановления или целевой области. С этой целью, выполняется анализ подобия между различными идентифицированными исходными областями, и когда исходный фрагмент является
40 аналогичным другим исходным фрагментам с подобием выше порогового значения, то этот исходный фрагмент может быть отброшен из набора потенциальных исходных фрагментов, поскольку он имеет высокую корреляцию с другими исходными фрагментами. Кроме того, в качестве типа стабилизации выбора фрагментов, предпочтительно поддерживать порядок фрагментов от предыдущего кадра, если ни
45 один из исходных фрагментов в текущем кадре не коррелируется (лучше данного порогового значения) с целевыми фрагментами в текущем кадре.

Дополнительный аспект основан на таких выявленных сведениях, что повышенное качество и уменьшенная скорость передачи битов, в частности, для сигналов,

содержащих переходные части, поскольку они возникают очень часто в аудиосигналах, получается за счет комбинирования технологии временного формирования шума (TNS) или временного формирования фрагментов (TTS) с восстановлением высоких частот. TNS/TTS-обработка на стороне кодера, реализуемая посредством прогнозирования по частоте, восстанавливает временную огибающую аудиосигнала. В зависимости от реализации, т.е. когда фильтр для временного формирования шума определяется не только в частотном диапазоне, покрывающем исходный частотный диапазон, но также и в целевом частотном диапазоне, который должен быть восстановлен в декодере повторного формирования частоты, временная огибающая применяется не только к базовому аудиосигналу до начальной частоты заполнения интервалов отсутствия сигнала, но временная огибающая также применяется к спектральным диапазонам восстановленных вторых спектральных частей. Таким образом, опережающие эхо или запаздывающие эхо, которые должны возникать без временного формирования фрагментов, уменьшаются или исключаются. Это достигается посредством применения обратного прогнозирования по частоте не только в базовом частотном диапазоне вплоть до определенной начальной частоты заполнения интервалов отсутствия сигнала, но также и в частотном диапазоне выше базового частотного диапазона. С этой целью, повторное формирование частоты или формирование частотных фрагментов выполняется на стороне декодера до применения прогнозирования по частоте. Тем не менее, прогнозирование по частоте может применяться либо до, либо после формирования спектральной огибающей в зависимости от того, вычисление информации энергии выполнено для остаточных спектральных значений после фильтрации или для (полных) спектральных значений перед формированием огибающей.

TTS-обработка для одного или более частотных фрагментов дополнительно устанавливает непрерывность корреляции между исходным диапазоном и диапазоном восстановления либо в двух смежных диапазонах восстановления или частотных фрагментах.

В реализации, предпочтительно использовать комплексную TNS/TTS-фильтрацию. В силу этого, не допускаются артефакты (временного) наложения спектров критически дискретизированного действительного представления, такого как MDCT. Комплексный TNS-фильтр может вычисляться на стороне кодера посредством применения не только модифицированного дискретного косинусного преобразования, но помимо этого, также и модифицированного дискретного синусного преобразования, чтобы получать комплексное модифицированное преобразование. Тем не менее, передаются только значения модифицированного дискретного косинусного преобразования, т.е. действительная часть комплексного преобразования. Тем не менее, на стороне декодера, можно оценивать мнимую часть преобразования с использованием MDCT-спектров предшествующих или последующих кадров, так что на стороне декодера комплексный фильтр может снова применяться при обратном прогнозировании по частоте и, в частности, при прогнозировании по границе между исходным диапазоном и диапазоном восстановления, а также по границе между смежными по частоте частотными фрагментами в диапазоне восстановления.

Изобретаемая система кодирования аудио эффективно кодирует произвольные аудиосигналы в широком диапазоне скоростей передачи битов. При этом, что для высоких скоростей передачи битов изобретаемая система стремится к прозрачности, для низких скоростей передачи битов минимизируется перцепционное раздражение. Следовательно, основная доля доступной скорости передачи битов используется для того, чтобы кодировать на основе формы сигналов только перцепционно наиболее

релевантную структуру сигнала в кодере, и результирующие интервалы отсутствия сигнала в спектре заполняются в декодере содержимым сигнала, который примерно аппроксимирует исходный спектр. Очень ограниченный битовый бюджет расходуется для того, чтобы управлять так называемым интеллектуальным заполнением интервалов (IGF) в спектре на основе параметров посредством выделенной вспомогательной информации, передаваемой из кодера в декодер.

Далее описываются предпочтительные варианты осуществления настоящего изобретения со ссылками на прилагаемые чертежи, на которых:

Фиг. 1a иллюстрирует устройство для кодирования аудиосигнала;

Фиг. 1b иллюстрирует декодер для декодирования кодированного аудиосигнала, совпадающий с кодером по фиг. 1a;

Фиг. 2a иллюстрирует предпочтительную реализацию декодера;

Фиг. 2b иллюстрирует предпочтительную реализацию кодера;

Фиг. 3a иллюстрирует схематичное представление спектра, сформированного посредством декодера в спектральной области по фиг. 1b;

Фиг. 3b иллюстрирует таблицу, указывающую взаимосвязь между коэффициентами масштабирования для полос частот коэффициентов масштабирования и энергиями для полос частот восстановления и информацией заполнения шумом для полосы частот заполнения шумом;

Фиг. 4a иллюстрирует функциональность кодера в спектральной области для применения выбора спектральных частей к первому и второму наборам спектральных частей;

Фиг. 4b иллюстрирует реализацию функциональности по фиг. 4a;

Фиг. 5a иллюстрирует функциональность MDCT-кодера;

Фиг. 5b иллюстрирует функциональность декодера с MDCT-технологией;

Фиг. 5c иллюстрирует реализацию модуля повторного формирования частоты;

Фиг. 6a иллюстрирует аудиокодер с функциональностью временного формирования шума/временного формирования фрагментов;

Фиг. 6b иллюстрирует декодер с технологией временного формирования шума/временного формирования фрагментов;

Фиг. 6c иллюстрирует дополнительную функциональность для функциональности временного формирования шума/временного формирования фрагментов с другим порядком спектрального прогнозного фильтра и спектрального формирователя;

Фиг. 7a иллюстрирует реализацию функциональности временного формирования фрагментов (TTS);

Фиг. 7b иллюстрирует реализацию декодера, совпадающую с реализацией кодера по фиг. 7a;

Фиг. 7c иллюстрирует спектрограмму исходного сигнала и расширенного сигнала без TTS;

Фиг. 7d иллюстрирует частотное представление, иллюстрирующее соответствие между частотами интеллектуального заполнения интервалов и энергиями временного формирования фрагментов;

Фиг. 7e иллюстрирует спектрограмму исходного сигнала и расширенного сигнала с TTS;

Фиг. 8a иллюстрирует двухканальный декодер с повторным формированием частоты;

Фиг. 8b иллюстрирует таблицу, иллюстрирующую различные комбинации представлений и исходных/целевых диапазонов;

Фиг. 8c иллюстрирует блок-схему последовательности операций способа,

иллюстрирующую функциональность двухканального декодера с повторным формированием частоты по фиг. 8а;

Фиг. 8d иллюстрирует более подробную реализацию декодера по фиг. 8а;

Фиг. 8е иллюстрирует реализацию кодера для двухканальной обработки, которая
5 должна декодироваться посредством декодера по фиг. 8а:

Фиг. 9а иллюстрирует декодер с технологией повторного формирования частоты с использованием значений энергии для частотного диапазона повторного формирования;

Фиг. 9b иллюстрирует более подробную реализацию модуля повторного формирования частоты по фиг. 9а;

10 Фиг. 9с иллюстрирует схематический вид, иллюстрирующий функциональность по фиг. 9b;

Фиг. 9d иллюстрирует дополнительную реализацию декодера по фиг. 9а;

Фиг. 10а иллюстрирует блок-схему кодера, совпадающего с декодером по фиг. 9а;

Фиг. 10b иллюстрирует блок-схему для иллюстрации дополнительной
15 функциональности модуля вычисления параметров по фиг. 10а;

Фиг. 10с иллюстрирует блок-схему, иллюстрирующую дополнительную функциональность параметрического модуля вычисления по фиг. 10а;

Фиг. 10d иллюстрирует блок-схему, иллюстрирующую дополнительную функциональность параметрического модуля вычисления по фиг. 10а;

20 Фиг. 11а иллюстрирует дополнительный декодер, имеющий конкретный идентификатор исходного диапазона для операции заполнения спектральными фрагментами в декодере;

Фиг. 11b иллюстрирует дополнительную функциональность модуля повторного формирования частоты по фиг. 11а;

25 Фиг. 11с иллюстрирует кодер, используемый для взаимодействия с декодером на фиг. 11а;

Фиг. 11d иллюстрирует блок-схему реализации модуля вычисления параметров по фиг. 11с;

Фиг. 12а и 12b иллюстрируют упрощенные графические схемы частоты для
30 иллюстрации исходного диапазона и целевого диапазона;

Фиг. 12с иллюстрирует график примерной корреляции двух сигналов;

Фиг. 13а иллюстрирует кодер предшествующего уровня техники с расширением полосы пропускания; и

Фиг. 13b иллюстрирует декодер предшествующего уровня техники с расширением
35 полосы пропускания.

Фиг. 1а иллюстрирует устройство для кодирования аудиосигнала 99. Аудиосигнал 99 вводится во временно-спектральный преобразователь 100 для преобразования аудиосигнала, имеющего частоту дискретизации, в спектральное представление 101, выводимое посредством временно-спектрального преобразователя. Спектр 101 вводится
40 в спектральный анализатор 102 для анализа спектрального представления 101.

Спектральный анализатор 101 выполнен с возможностью определения первого набора первых спектральных частей 103, которые должны быть кодированы с первым спектральным разрешением, и другого второго набора вторых спектральных частей 105, которые должны быть кодированы со вторым спектральным разрешением. Второе спектральное разрешение меньше первого спектрального разрешения. Второй набор вторых спектральных частей 105 вводится в модуль 104 вычисления параметров или параметрический кодер для вычисления информации спектральной огибающей, имеющей второе спектральное разрешение. Кроме того, аудиокодер 106 в спектральной области

предоставляется для формирования первого кодированного представления 107 первого набора первых спектральных частей, имеющих первое спектральное разрешение. Кроме того, модуль 104 вычисления параметров/параметрический кодер выполнен с возможностью формирования второго кодированного представления 109 второго набора вторых спектральных частей. Первое кодированное представление 107 и второе кодированное представление 109 вводятся в мультиплексор потоков битов или формирователь 108 потоков битов, и блок 108, в завершение, выводит кодированный аудиосигнал для передачи или хранения на устройстве хранения данных.

Типично, первая спектральная часть, к примеру, 306 по фиг. 3а, окружена посредством двух вторых спектральных частей, к примеру, 307а, 307b. Дело обстоит не так в HE-AAC, в котором частотный диапазон базового кодера имеет ограниченную полосу частот.

Фиг. 1b иллюстрирует декодер, совпадающий с кодером по фиг. 1а. Первое кодированное представление 107 вводится в аудиодекодер 112 в спектральной области для формирования первого декодированного представления первого набора первых спектральных частей, причем декодированное представление имеет первое спектральное разрешение. Кроме того, второе кодированное представление 109 вводится в параметрический декодер 114 для формирования второго декодированного представления второго набора вторых спектральных частей, имеющих второе спектральное разрешение ниже первого спектрального разрешения.

Декодер дополнительно содержит модуль 116 повторного формирования частоты для повторного формирования восстановленной второй спектральной части, имеющей первое спектральное разрешение, с использованием первой спектральной части. Модуль 116 повторного формирования частоты выполняет операцию заполнения фрагментами, т.е. использует фрагмент или часть первого набора первых спектральных частей и копирует этот первый набор первых спектральных частей в диапазон восстановления или полосу частот восстановления, имеющую вторую спектральную часть, и типично выполняет формирование спектральной огибающей или другую операцию, как указано посредством декодированного второго представления, выводимого посредством параметрического декодера 114, т.е. посредством использования информации относительно второго набора вторых спектральных частей. Декодированный первый набор первых спектральных частей и восстановленный второй набор спектральных частей, как указано на выходе модуля 116 повторного формирования частоты на линии 117, вводятся в спектрально-временной преобразователь 118, выполненный с возможностью преобразования первого декодированного представления и восстановленной второй спектральной части во временное представление 119, причем временное представление имеет определенную высокую частоту дискретизации.

Фиг. 2b иллюстрирует реализацию кодера по фиг. 1а. Входной аудиосигнал 99 вводится в гребенку 220 аналитических фильтров, соответствующую временно-спектральному преобразователю 100 по фиг. 1а. Затем операция временного формирования шума выполняется в TNS-блоке 222. Следовательно, ввод в спектральный анализатор 102 по фиг. 1а, соответствующий блочной тональной маске 226 по фиг. 2b, может представлять собой либо полные спектральные значения, когда не применяется операция временного формирования шума/временного формирования фрагментов, либо может представлять собой остаточные спектральные значения, когда применяется TNS-операция, как проиллюстрировано на фиг. 2b, блок 222. Для двухканальных сигналов или многоканальных сигналов, дополнительно может выполняться объединенное канальное кодирование 228, так что кодер 106 в спектральной области

по фиг. 1а может содержать блок 228 объединенного канального кодирования. Кроме того, предоставляется энтропийный кодер 232 для выполнения сжатия данных без потерь, который также является частью кодера 106 в спектральной области по фиг. 1а.

Спектральный анализатор/тональная маска 226 разделяет вывод TNS-блока 222 на полосу базовых частот и тональные компоненты, соответствующие первому набору первых спектральных частей 103, и остаточные компоненты, соответствующие второму набору вторых спектральных частей 105 по фиг. 1а. Блок 224, указываемый в качестве кодирования с извлечением IGF-параметров, соответствует параметрическому кодеру 104 по фиг. 1а, а мультиплексор 230 потоков битов соответствует мультиплексору 108 потоков битов по фиг. 1а.

Предпочтительно, гребенка аналитических фильтров 222 реализована как MDCT (гребенка фильтров модифицированного дискретного косинусного преобразования), и MDCT используется для того, чтобы преобразовывать сигнал 99 в частотно-временную область с помощью модифицированного дискретного косинусного преобразования, выступающего в качестве инструментального средства частотного анализа.

Спектральный анализатор 226 предпочтительно применяет маску тональности. Эта стадия оценки масок тональности используется для того, чтобы отделять тональные компоненты от шумоподобных компонентов в сигнале. Это дает возможность базовому кодеру 228 кодировать все тональные компоненты с помощью психоакустического модуля. Стадия оценки масок тональности может реализовываться множеством различных способов и предпочтительно реализуется аналогично по функциональности стадии оценки синусоидальных дорожек, используемой при синусоидальном и шумовом моделировании для кодирования речи/аудио [8, 9] или в аудиокодере на основе HILN-модели, описанном в [10]. Предпочтительно, используется реализация, которую просто реализовывать без необходимости поддерживать траектории процесса "рождения и гибели", но также может использоваться любой другой детектор тональности или шума.

IGF-модуль вычисляет подобие, которое существует между исходной областью и целевой областью. Целевая область должна быть представлена посредством спектра из исходной области. Мера подобия между исходными и целевыми областями задается с использованием подхода на основе взаимной корреляции. Целевая область разбивается на *nTar* неперекрывающихся частотных фрагментов. Для каждого фрагмента в целевой области, *nSrc* исходных фрагментов создаются из фиксированной начальной частоты. Эти исходные фрагменты перекрываются на коэффициент между 0 и 1, где 0 означает 0%-ое перекрытие, а 1 означает 100%-ое перекрытие. Каждый из этих исходных фрагментов коррелирован с целевым фрагментом с различными запаздываниями с тем, чтобы находить исходный фрагмент, который наилучше совпадает с целевым фрагментом. Наилучше совпадающий номер фрагмента сохраняется в *tileNum[idx_tar]*, запаздывание, при котором он лучше всего коррелируется с целью, сохраняется в *xcorr_lag[idx_tar][idx_src]*, а знак корреляции сохраняется в *xcorr_sign[idx_tar][idx_src]*. В случае если корреляция является сильно отрицательной, исходный фрагмент должен быть умножен на -1 перед процессом заполнения фрагментами в декодере. IGF-модуль также отслеживает непerezапись тональных компонентов в спектре, поскольку тональные компоненты сохраняются с использованием маски тональности. Энергетический параметр для каждой полосы частот используется для того, чтобы сохранять энергию целевой области, позволяя точно восстанавливать спектр.

Этот способ имеет определенные преимущества по сравнению с классическим SBR [1] в том, что гармоническая сетка многотонального сигнала сохраняется посредством

базового кодера, тогда как только интервалы отсутствия сигнала между синусоидами заполнены наилучше совпадающим "шумом определенной формы" из исходной области. Другое преимущество этой системы по сравнению с ASR (точной спектральной заменой) [2-4] заключается в отсутствии стадии синтеза сигналов, которая создает важные части сигнала в декодере. Вместо этого, выполнение этой задачи возлагается на базовый кодер, обеспечивая сохранение важных компонентов спектра. Другое преимущество предложенной системы заключается в непрерывной масштабируемости, которую предлагают признаки. Использование только `tileNum[idx_tar]` и `xcorr_lag = 0` для каждого фрагмента называется "общим совпадением степени детализации" и может использоваться для низких скоростей передачи битов, тогда как использование переменной `xcorr_lag` для каждого фрагмента позволяет обеспечивать лучшее совпадение целевых и исходных спектров.

Помимо этого, предложена технология стабилизации выбора фрагментов, которая удаляет артефакты в частотной области, такие как расстроение и музыкальный шум.

В случае пар стереоканалов применяется дополнительная объединенная стереообработка. Это необходимо, поскольку для определенного целевого диапазона сигнал может высокоррелированный панорамированный источник звука. В случае если исходные области, выбранные для этой конкретной области, не имеют хорошей корреляции, хотя энергии совпадают для целевых областей, пространственное изображение может ухудшаться некоррелированных исходных областей. Кодер анализирует каждую энергетическую полосу частот целевой области, типично выполняя взаимную корреляцию спектральных значений, и если определенное пороговое значение превышает, задает объединенный флаг для этой энергетической полосы частот. В декодере, энергетические полосы частот левого и правого канала обрабатываются по отдельности, если этот флаг объединенного стерео не задан. В случае если флаг объединенного стерео задается, как энергии, так и наложение выполняются в объединенной стереобласти. Объединенная стереоинформация для IGF-областей передается в служебных сигналах, аналогично объединенной стереоинформации для базового кодирования, включающей в себя флаг, указывающий, в случае прогнозирования, то, представляет собой направление прогнозирования от низведения до остатка или наоборот.

Энергии могут вычисляться из передаваемых энергий в L/R-области.

$$midNrg[k] = leftNrg[k] + rightNrg[k];$$

$$sideNrg[k] = leftNrg[k] - rightNrg[k];$$

где k является частотным индексом в области преобразования.

Другое решение состоит в том, чтобы вычислять и передавать энергии непосредственно в объединенной стереобласти для полос частот, в которых объединенное стерео является активным, так что дополнительное преобразование энергии не требуется на стороне декодера.

Исходные фрагменты всегда создаются согласно матрице среднего/бокового каналов:

$$midTile[k] = 0.5 \cdot (leftTile[k] - rightTile[k])$$

$$sideTile[k] = 0.5 \cdot (leftTile[k] + rightTile[k])$$

Энергетическое регулирование:

$$midTile[k] = midTile[k] * midNrg[k];$$

$$sideTile[k] = sideTile[k] * sideNrg[k];$$

Объединенное стерео->LR-преобразование:

Если дополнительные параметры прогнозирования не кодируются:

$$\mathbf{leftTile[k] = midTile[k] - sideTile[k]}$$

$$\mathbf{rightTile[k] = midTile[k] - sideTile[k]}$$

Если дополнительный параметр прогнозирования кодируется, и если передаваемое в служебных сигналах направление представляет собой от среднего к боковому:

$$sideTile[k] = sideTile[k] - predictionCoeff \cdot midTile[k]$$

$$leftTile[k] = midTile[k] + sideTile[k]$$

$$rightTile[k] = midTile[k] - sideTile[k]$$

Если передаваемое в служебных сигналах направление представляет собой от бокового к среднему:

$$midTile1[k] = midTile[k] - predictionCoeff \cdot sideTile[k]$$

$$leftTile[k] = midTile1[k] - sideTile[k]$$

$$rightTile[k] = midTile1[k] + sideTile[k]$$

Эта обработка обеспечивает то, что из фрагментов, используемых для повторного формирования высокоррелированных целевых областей и панорамированных целевых областей, результирующие левый и правый каналы по-прежнему представляют коррелированный и панорамированный источник звука, даже если исходные области не коррелированы, сохраняя стереоизображение для таких областей.

Другими словами, в потоке битов передаются флаги объединенного стерео, которые указывают то, должен использоваться L/R или M/S в качестве примера для общего объединенного стереокодирования. В декодере, во-первых декодируется базовый сигнал, как указано посредством флагов объединенного стерео для полос базовых частот. Во-вторых, базовый сигнал сохраняется в обоих L/R- и M/S-представлениях. Для заполнения IGF-фрагментами, представление исходных фрагментов выбрано таким образом, что оно соответствует представлению целевых фрагментов, как указано посредством объединенной стереоинформации для полос IGF-частот.

Временное формирование шума (TNS) является стандартной технологией и частью AAC [11-13]. TNS может рассматриваться как расширение базовой схемы перцепционного кодера, вставляющее необязательный этап обработки между гребенкой фильтров и стадией квантования. Основная задача модуля TNS состоит в том, чтобы скрывать сформированный шум квантования в области временного маскирования переходных сигналов, в силу чего это приводит к более эффективной схеме кодирования. Во-первых, TNS вычисляет набор коэффициентов прогнозирования с использованием "прямого прогнозирования" в области преобразования, например, MDCT. Эти коэффициенты затем используются для сглаживания временной огибающей сигнала. Поскольку квантование затрагивает TNS-фильтрованный спектр, также шум квантования является временно плоским. Посредством применения обратной TNS-фильтрации на стороне декодера, шуму квантования придается определенная форма согласно временной огибающей TNS-фильтра, и следовательно, шум квантования маскируется посредством переходной части.

IGF основано на MDCT-представлении. Для эффективного кодирования предпочтительно должны использоваться длинные блоки приблизительно в 20 мс. Если сигнал в таком длинном блоке содержит переходные части, слышимые опережающие и запаздывающие эхо возникают в полосах IGF-спектра вследствие заполнения

фрагментами. Фиг. 7с показывает типичный эффект опережающего эхо перед началом переходной части вследствие IGF. Слева показана спектрограмма исходного сигнала, а справа показана спектрограмма сигнала с расширенной полосой пропускания без TNS-фильтрации.

Этот эффект опережающего эхо уменьшается посредством использования TNS в IGF-контексте. Здесь, TNS используется в качестве инструментального средства временного формирования фрагментов (TTS), поскольку повторное формирование спектра в декодере выполняется для остаточного TNS-сигнала. Требуемые коэффициенты TTS-прогнозирования вычисляются и применяются с использованием полного спектра на стороне кодера, как обычно. Начальные и конечные TNS/TTS-частоты не затрагиваются посредством начальной IGF-частоты $f_{IGFstart}$ инструментального IGF-средства. По сравнению с унаследованным TNS, конечная TTS-частота увеличивается до конечной частоты инструментального IGF-средства, которая выше $f_{IGFstart}$. На стороне декодера, TNS/TTS-коэффициенты применяются к полному спектру снова, т.е. к базовому спектру плюс повторно сформированный спектр плюс тональные компоненты из карты тональности (см. фиг. 7e). Применение TTS необходимо для того, чтобы снова формировать временную огибающую повторно сформированного спектра таким образом, что она совпадает с огибающей исходного сигнала. Таким образом, уменьшаются показанные опережающие эхо. Помимо этого, оно по-прежнему придает определенную форму шуму квантования в сигнале ниже $f_{IGFstart}$, что является стандартным для TNS.

В унаследованных декодерах, спектральное наложение на аудиосигнал нарушает спектральную корреляцию на границах наложений и в силу этого искажает временную огибающую аудиосигнала посредством введения дисперсии. Следовательно, другое преимущество выполнения заполнения IGF-фрагментами для остаточного сигнала заключается в том, что после применения формирующего фильтра границы фрагментов прозрачно коррелированы, приводя к более достоверному временному воспроизведению сигнала.

В изобретаемом кодере, спектр, подвергнутый TNS/TTS-фильтрации, обработке маски тональности и оценке IGF-параметров, не имеет сигналов выше начальной IGF-частоты за исключением тональных компонентов. Далее этот разреженный спектр кодируется посредством базового кодера с использованием принципов арифметического кодирования и прогнозирующего кодирования. Эти кодированные компоненты вместе с сигнальными битами формируют поток битов аудио.

Фиг. 2а иллюстрирует соответствующую реализацию декодера. Поток битов на фиг. 2а, соответствующий кодированному аудиосигналу, вводится в демультимплексор/декодер, который должен быть соединен, относительно фиг. 1b, с блоками 112 и 114. Демультимплексор потоков битов разделяет входной аудиосигнал на первое кодированное представление 107 по фиг. 1b и второе кодированное представление 109 по фиг. 1b. Первое кодированное представление, имеющее первый набор первых спектральных частей, вводится в блок 204 объединенного канального декодирования, соответствующий декодеру 112 в спектральной области по фиг. 1b. Второе кодированное представление вводится в параметрический декодер 114, не проиллюстрированный на фиг. 2а, а затем вводится в IGF-блок 202, соответствующий модулю 116 повторного формирования частоты по фиг. 1b. Первый набор первых спектральных частей, требуемых для повторного формирования частоты, вводится в IGF-блок 202 через линию 203. Кроме того, после объединенного канального декодирования 204,

конкретное базовое декодирование применяется в блоке 206 применения тональной маски, так что вывод тональной маски 206 соответствует выводу декодера 112 в спектральной области. Затем выполняется комбинирование посредством модуля 208 комбинирования, т.е. компоновка кадров, причем вывод модуля 208 комбинирования в данный момент имеет полнодиапазонный спектр, но при этом в TNS/TTS-фильтрованной области. Затем в блоке 210 обратная TNS/TTS-операция выполняется с использованием информации TNS/TTS-фильтра, предоставленной через линию 109, т.е. вспомогательная TTS-информация предпочтительно включена в первое кодированное представление, сформированное посредством кодера 106 в спектральной области, который, например, может представлять собой простой базовый AAC- или USAC-кодер, либо также может быть включена во второе кодированное представление. На выходе блока 210 предоставляется полный спектр до максимальной частоты, которая представляет собой полнодиапазонную частоту, заданную посредством частоты дискретизации исходного входного сигнала. Затем спектрально-временное преобразование выполняется в гребенке 212 синтезирующих фильтров с тем, чтобы, в завершение, получать выходной аудиосигнал.

Фиг. 3а иллюстрирует схематичное представление спектра. Спектр подразделяется на полосы SCB частот коэффициентов масштабирования, причем предусмотрено семь полос SCB1-SCB7 частот коэффициентов масштабирования в проиллюстрированном примере по фиг. 3а. Полосы частот коэффициентов масштабирования могут представлять собой полосы частот AAC-коэффициентов масштабирования, которые задаются в AAC-стандарте и имеют возрастающую полосу пропускания до верхних частот, как схематично проиллюстрировано на фиг. 3а. Предпочтительно выполнять интеллектуальное заполнение интервалов не с самого начала спектра, т.е. на низких частотах, а начинать IGF-операцию на начальной IGF-частоте, проиллюстрированной на 309. Следовательно, полоса базовых частот идет от наименьшей частоты до начальной IGF-частоты. Выше начальной IGF-частоты, применяется спектральный анализ для того, чтобы отделять спектральные компоненты 304, 305, 306, 307 высокого разрешения (первый набор первых спектральных частей) от компонентов низкого разрешения, представленных посредством второго набора вторых спектральных частей. Фиг. 3а иллюстрирует спектр, который примерно вводится в кодер 106 в спектральной области или объединенный канальный кодер 228, т.е. базовый кодер работает в полном диапазоне, но кодирует существенное число нулевых спектральных значений, т.е. эти нулевые спектральные значения квантованы до нуля или задаются равными нулю до квантования или после квантования. В любом случае, базовый кодер работает в полном диапазоне, т.е. как если спектр является таким, как проиллюстрировано, т.е. базовый декодер не обязательно должен знать об интеллектуальном заполнении интервалов или кодировании второго набора вторых спектральных частей с более низким спектральным разрешением.

Предпочтительно, высокое разрешение задается посредством полинейного кодирования спектральных линий, таких как MDCT-линии, тогда как второе разрешение или низкое разрешение задается, например, посредством вычисления только одного спектрального значения в расчете на полосу частот коэффициентов масштабирования, при этом полоса частот коэффициентов масштабирования покрывает несколько частотных линий. Таким образом, второе низкое разрешение, относительно спектрального разрешения, гораздо ниже первого или высокого разрешения, заданного посредством полинейного кодирования, типично применяемого посредством базового кодера, к примеру, базового AAC- или USAC-кодера.

Относительно вычисления коэффициентов масштабирования или энергии, ситуация проиллюстрирована на фиг. 3b. Вследствие того факта, что кодер представляет собой базовый кодер, и вследствие того факта, что могут (но не обязательно должны) быть предусмотрены компоненты первого набора спектральных частей в каждой полосе частот, базовый кодер вычисляет коэффициент масштабирования для каждой полосы частот не только в базовом диапазоне ниже начальной IGF-частоты 309, но также и выше начальной IGF-частоты вплоть до максимальной частоты $f_{IGFstop}$, которая меньше или равна половине частоты дискретизации, т.е. $f_s/2$. Таким образом, кодированные тональные части 302, 304, 305, 306, 307 по фиг. 3a, в этом варианте осуществления, вместе с коэффициентами SCB1-SCB7 масштабирования соответствуют спектральным данным высокого разрешения. Спектральные данные низкого разрешения вычисляются, начиная с начальной IGF-частоты, и соответствуют значениям E1, E2, E3, E4 информации энергии, которые передаются вместе с коэффициентами SF4-SF7 масштабирования.

В частности, когда базовый кодер применяется в состоянии с низкой скоростью передачи битов, помимо этого, может применяться дополнительная операция заполнения шумом в полосе базовых частот, т.е. ниже по частоте по сравнению с начальной IGF-частотой, т.е. в полосах SCB1-SCB3 частот коэффициентов масштабирования. При заполнении шумом, существует несколько смежных спектральных линий, которые квантованы до нуля. На стороне декодера, эти квантованные до нуля спектральные значения повторно синтезируются, и повторно синтезированные спектральные значения регулируются по абсолютной величине с использованием энергии заполнения шумом, к примеру, NF2, проиллюстрированной на 308 на фиг. 3b. Энергия заполнения шумом, которая может предоставляться в абсолютном выражении или в относительном выражении, в частности, относительно коэффициента масштабирования, что характерно для USAC, соответствует энергии набора спектральных значений, квантованных до нуля. Эти спектральные линии заполнения шумом также могут рассматриваться в качестве третьего набора третьих спектральных частей, которые повторно формируются посредством прямого синтеза заполнения шумом без IGF-операции на основе повторного формирования частоты с использованием частотных фрагментов из других частот, для восстановления частотных фрагментов с использованием спектральных значений из исходного диапазона и информации E1, E2, E3, E4 энергии.

Предпочтительно, полосы частот, для которых вычисляется информация энергии, совпадают с полосами частот коэффициентов масштабирования. В других вариантах осуществления, применяется группировка значения информации энергии, так что, например, для полос 4 и 5 частот коэффициентов масштабирования, передается только одно значение информации энергии, но даже в этом варианте осуществления границы сгруппированных полос частот восстановления совпадают с границами полос частот коэффициентов масштабирования. Если применяются различные разделения полос частот, то определенные повторные вычисления или вычисления синхронизации могут применяться, и это может быть целесообразным в зависимости от определенной реализации.

Предпочтительно, кодер 106 в спектральной области по фиг. 1a представляет собой психоакустически регулируемый кодер, как проиллюстрировано на фиг. 4a. Типично, как, например, проиллюстрировано в стандарте MPEG2/4 AAC или в стандарте MPEG1/2 Layer 3, аудиосигнал, который должен быть кодирован, после преобразования в спектральный диапазон (401 на фиг. 4a) перенаправляется в модуль 400 вычисления коэффициентов масштабирования. Модуль вычисления коэффициентов масштабирования управляется посредством психоакустической модели, дополнительно

принимавшей подлежащий квантованию аудиосигнал или принимающей, согласно стандарту MPEG1/2 Layer 3 или MPEG AAC, комплексное спектральное представление аудиосигнала. Психоакустическая модель вычисляет, для каждой полосы частот коэффициентов масштабирования, коэффициент масштабирования, представляющий психоакустическое пороговое значение. Кроме того, коэффициенты масштабирования затем, посредством взаимодействия известных внутренних и внешних итерационных циклов либо посредством любой другой подходящей процедуры кодирования, регулируются таким образом, что удовлетворяются определенные условия по скорости передачи битов. Далее подлежащие квантованию спектральные значения, с одной стороны, и вычисленные коэффициенты масштабирования, с другой стороны, вводятся в процессор 404 квантователя. В простом алгоритме работы аудиокодера, подлежащие квантованию спектральные значения взвешиваются посредством коэффициентов масштабирования, и взвешенные спектральные значения затем вводятся в квантователь с фиксированным шагом, типично имеющий функциональность сжатия до диапазонов верхних амплитуд. Затем на выходе процессора квантователя, существуют индексы квантования, которые затем передаются в энтропийный кодер, типично имеющий конкретное и очень эффективное кодирование для набора нулевых индексов квантования для смежных значений частоты или, как они также называются в данной области техники, для "серий" нулевых значений.

Тем не менее, в аудиокодере по фиг. 1a, процессор квантователя типично принимает информацию относительно вторых спектральных частей из спектрального анализатора. Таким образом, процессор 404 квантователя удостоверяется, что на выходе процессора 404 квантователя вторые спектральные части, идентифицированные посредством спектрального анализатора 102, являются нулевыми или имеют представление, подтвержденное посредством кодера или декодера как нулевое представление, которое может быть очень эффективно кодировано, в частности, когда существуют "серии" нулевых значений в спектре.

Фиг. 4b иллюстрирует реализацию процессора квантователя. Спектральные MDCT-значения могут вводиться в блок 410 обнуления. Затем вторые спектральные части уже задаются равными нулю до того, как выполняется взвешивание посредством коэффициентов масштабирования в блоке 412. В дополнительной реализации, блок 410 не предоставляется, а взаимодействие для обнуления выполняется в блоке 418 после взвешивающего блока 412. В еще одной дополнительной реализации, операция обнуления также может выполняться в блоке 422 обнуления после квантования в блоке 420 квантователя. В этой реализации, блоки 410 и 418 не должны присутствовать. Обычно по меньшей мере один из блоков 410, 418, 422 предоставляется в зависимости от конкретной реализации.

Затем на выходе блока 422 получается квантованный спектр согласно тому, что проиллюстрировано на фиг. 3a. Этот квантованный спектр затем вводится в энтропийный кодер, такой как 232 на фиг. 2b, который может представлять собой кодер Хаффмана или арифметический кодер, например, заданный в USAC-стандарте.

Блоки 410, 418, 422 обнуления, которые предоставляются альтернативно друг другу или параллельно, управляются посредством спектрального анализатора 424.

Спектральный анализатор предпочтительно содержит любую реализацию известного детектора тональности или содержит любой другой тип детектора, выполненного с возможностью разделения спектра на компоненты, которые должны быть кодированы с высоким разрешением, и компоненты, которые должны быть кодированы с низким разрешением. Другие такие алгоритмы, реализованные в спектральном анализаторе,

могут представлять собой детектор речевой активности, детектор шума, детектор речи или любой другой детектор, определяющий, в зависимости от спектральной информации или ассоциированных метаданных, требования по разрешению для различных спектральных частей.

5 Фиг. 5а иллюстрирует предпочтительную реализацию временно-спектрального преобразователя 100 по фиг. 1а, например, реализованного в AAC или USAC. Временно-спектральный преобразователь 100 содержит модуль 502 кодирования со взвешиванием, управляемый посредством детектора 504 переходных частей. Когда детектор 504 переходных частей обнаруживает переходную часть, в таком случае переключение с
10 длинных окон кодирования со взвешиванием на короткие окна кодирования со взвешиванием передается в служебных сигналах в модуль кодирования со взвешиванием. Модуль 502 кодирования со взвешиванием затем вычисляет, для перекрывающихся блоков, кодированные со взвешиванием кадры, причем каждый кодированный со взвешиванием кадр типично имеет два N значений, к примеру, 2048 значений. После
15 этого выполняется преобразование в модуле 506 блочного преобразования, и этот модуль блочного преобразования типично дополнительно предоставляет прореживание, так что комбинированное прореживание/преобразование выполняется для того, чтобы получать спектральный кадр с N значениями, к примеру, спектральными MDCT-значениями. Таким образом, для работы в периоды длинных окон кодирования со
20 взвешиванием, кадр на входе блока 506 содержит два N значений, к примеру, 2048 значений, и спектральный кадр в таком случае имеет 1024 значения. Тем не менее, затем выполняется переключение на короткие блоки, когда выполняются восемь коротких блоков, при этом каждый короткий блок имеет 1/8 от числа кодированных со взвешиванием значений во временной области по сравнению с длинным окном
25 кодирования со взвешиванием, и каждый спектральный блок имеет 1/8 от числа спектральных значений по сравнению с длинным блоком. Таким образом, когда это прореживание комбинировано с операцией 50%-ого перекрытия модуля кодирования со взвешиванием, спектр является критически дискретизированной версией аудиосигнала
99 временной области.

30 Далее следует обратиться к фиг. 5b, иллюстрирующему конкретную реализацию модуля 116 повторного формирования частоты и спектрально-временного преобразователя 118 по фиг. 1b или комбинированной работы блоков 208, 212 по фиг. 2а. На фиг. 5b, рассматривается конкретная полоса частот восстановления, к примеру, полоса 6 частот коэффициентов масштабирования по фиг. 3а. Первая спектральная
35 часть в этой полосе частот восстановления, т.е. первая спектральная часть 306 по фиг. 3а вводится в блок 510 компоновки/регулирования кадров. Кроме того, восстановленная вторая спектральная часть для полосы 6 частот коэффициентов масштабирования также вводится в модуль 510 компоновки/регулирования кадров. Кроме того, информация энергии, к примеру, E3 по фиг. 3b для полосы 6 частот коэффициентов
40 масштабирования также вводится в блок 510. Восстановленная вторая спектральная часть в полосе частот восстановления уже сформирована посредством заполнения частотными фрагментами с использованием исходного диапазона, и полоса частот восстановления затем соответствует целевому диапазону. Далее, выполняется энергетическое регулирование кадра, чтобы затем в итоге получать полный
45 восстановленный кадр, имеющий N значений, например, получаемый на выходе модуля 208 комбинирования по фиг. 2а. После этого в блоке 512 обратное блочное преобразование/интерполяция выполняется для того, чтобы получать 248 значений во временной области, например, для 124 спектральных значений на входе блока 512.

Затем операция синтезирующего кодирования со взвешиванием выполняется в блоке 514, который снова управляется посредством индикатора длинного окна кодирования со взвешиванием/короткого окна кодирования со взвешиванием, передаваемого в качестве вспомогательной информации в кодированном аудиосигнале. После этого в блоке 516, выполняется операция суммирования с перекрытием с предыдущим временным кадром. Предпочтительно, MDCT применяет 50%-ое перекрытие, так что для каждого нового временного кадра с $2N$ значений, в итоге выводятся N значений во временной области. 50%-ое перекрытие в большой степени предпочитается вследствие того факта, что оно предоставляет критическую дискретизацию и непрерывное разделение от одного кадра до следующего кадра вследствие операции суммирования с перекрытием в блоке 516.

Как проиллюстрировано на 301 на фиг. 3а, операция заполнения шумом дополнительно может применяться не только ниже начальной IGF-частоты, но также и выше начальной IGF-частоты, к примеру, для рассмотренной полосы частот восстановления, совпадающей с полосой 6 частот коэффициентов масштабирования по фиг. 3а. Затем спектральные значения заполнения шумом также могут вводиться в модуль 510 компоновки/регулирования кадров, и регулирование спектральных значений заполнения шумом также может применяться в этом блоке, или спектральные значения заполнения шумом могут уже регулироваться с использованием энергии заполнения шумом до ввода в модуль 510 компоновки/регулирования кадров.

Предпочтительно, IGF-операция, т.е. операция заполнения частотными фрагментами с использованием спектральных значений из других частей, может применяться в полном спектре. Таким образом, операция заполнения спектральными фрагментами может не только применяться в полосе высоких частот выше начальной IGF-частоты, но также может применяться в полосе низких частот. Кроме того, заполнение шумом без заполнения частотными фрагментами также может применяться не только ниже начальной IGF-частоты, но также и выше начальной IGF-частоты. Тем не менее, обнаружено, что высокое качество и высокоэффективное кодирование аудио могут быть получены, когда операция заполнения шумом ограничена частотным диапазоном ниже начальной IGF-частоты, и когда операция заполнения частотными фрагментами ограничивается частотным диапазоном выше начальной IGF-частоты, как проиллюстрировано на фиг. 3а.

Предпочтительно, целевые фрагменты (ТТ) (имеющие частоты, большие начальной IGF-частоты), ограничены границами полос частот коэффициентов масштабирования полноскоростного кодера. Исходные фрагменты (ST), из которых извлекается информация, т.е. для частот ниже начальной IGF-частоты, не ограничены посредством границ полос частот коэффициентов масштабирования. Размер ST должен соответствовать размеру ассоциированного ТТ. Это проиллюстрировано с использованием следующего примера. ТТ[0] имеет длину в 10 элементов MDCT-выборки. Она точно соответствует длине двух последующих SCB (к примеру, 4+6). Далее все возможные ST, которые должны быть коррелированы с ТТ[0], также имеют длину в 10 элементов выборки. Второй целевой фрагмент ТТ[1], смежный с ТТ[0], имеет длину в 15 элементов выборки (SCB, имеющий длину 7+8). Далее ST для него имеют длину 15 элементов выборки, а не 10 элементов выборки, как для ТТ[0].

Если возникает такая ситуация, что невозможно находить ТТ для ST с длиной целевого фрагмента (когда, например, длина ТТ превышает доступный исходный диапазон), то корреляция не вычисляется, и исходный диапазон копируется определенное число раз в этот ТТ (копирование выполняется по одному элементу, так что частотная линия для

наименьшей частоты второй копии идет сразу – по частоте – после частотной линии для наибольшей частоты первой копии) до тех пор, пока целевой фрагмент (ТТ) не будет полностью заполнен.

В дальнейшем следует обратиться к фиг. 5с, иллюстрирующему дополнительный предпочтительный вариант осуществления модуля 116 повторного формирования частоты по фиг. 1b или IGF-блока 202 по фиг. 2а. Блок 522 представляет собой модуль формирования частотных фрагментов, принимающий не только идентификатор целевой полосы частот, но и дополнительно принимающий идентификатор исходной полосы частот. В качестве примера, на стороне кодера определено то, что полоса 3 частот коэффициентов масштабирования по фиг. 3а очень хорошо подходит для восстановления полосы 7 частот коэффициентов масштабирования. Таким образом, идентификатор исходной полосы частот должен составлять 2, а идентификатор целевой полосы частот должен составлять 7. На основе этой информации модуль 522 формирования частотных фрагментов применяет операцию заполнения фрагментов перезаписи или гармоник или любую другую операцию заполнения фрагментами, чтобы сформировать необработанную вторую часть спектральных компонентов 523. Необработанная вторая часть спектральных компонентов имеет частотное разрешение, идентичное частотному разрешению, включенному в первый набор первых спектральных частей.

Затем первая спектральная часть полосы частот восстановления, к примеру, 307 по фиг. 3а, вводится в модуль 524 компоновки кадров, и необработанная вторая часть 523 также вводится в модуль 524 компоновки кадров. Затем восстановленный кадр регулируется посредством модуля 526 регулирования с использованием коэффициента усиления для полосы частот восстановления, вычисленной посредством модуля 528 вычисления коэффициентов усиления. Тем не менее, важно, что первая спектральная часть в кадре не затрагивается посредством модуля 526 регулирования, а только необработанная вторая часть для кадра восстановления затрагивается посредством модуля 526 регулирования. С этой целью, модуль 528 вычисления коэффициентов усиления анализирует исходную полосу частот или необработанную вторую часть 523 и дополнительно анализирует первую спектральную часть в полосе частот восстановления, чтобы, в завершение, находить корректный коэффициент 527 усиления, так что энергия отрегулированного кадра, выводимого посредством модуля 526 регулирования, имеет энергию E4, когда рассматривается полоса 7 частот коэффициентов масштабирования.

В этом контексте, очень важно оценивать точность восстановления высоких частот настоящего изобретения по сравнению с HE-AAC. Это поясняется относительно полосы 7 частот коэффициентов масштабирования на фиг. 3а. Предполагается, что кодер предшествующего уровня техники, к примеру, проиллюстрированный на фиг. 13а, обнаруживает спектральную часть 307, которая должна кодироваться с высоким разрешением, в качестве "пропущенных гармоник". Затем энергия этого спектрального компонента передается вместе с информацией спектральной огибающей для полосы частот восстановления, к примеру, для полосы 7 частот коэффициентов масштабирования, в декодер. Далее декодер должен воссоздавать пропущенную гармонику. Тем не менее, спектральное значение, при котором пропущенная гармоника 307 восстанавливается посредством декодера предшествующего уровня техники по фиг. 13b, должно находиться в середине полосы частот 7 на частоте, указываемой посредством частоты 390 восстановления. Таким образом, настоящее изобретение избегает ошибки 391 по частоте, которая вводится посредством декодера предшествующего уровня техники по фиг. 13d.

В реализации, спектральный анализатор также реализован с возможностью вычислять подобия между первыми спектральными частями и вторыми спектральными частями и определять, на основе вычисленных подобий для второй спектральной части в диапазоне восстановления, первую спектральную часть, совпадающую со второй спектральной частью в максимально возможной степени. Затем в этой реализации с переменными исходными диапазонами/целевыми диапазонами параметрический кодер дополнительно вводит во второе кодированное представление информацию совпадения, указывающую для каждого целевого диапазона совпадающий исходный диапазон. На стороне декодера, эта информация затем используется посредством модуля 522 формирования частотных фрагментов по фиг. 5с, иллюстрирующего формирование необработанной второй части 523 на основе идентификатора исходной полосы частот и идентификатора целевой полосы частот.

Кроме того, как проиллюстрировано на фиг. 3а, спектральный анализатор выполнен с возможностью анализировать спектральное представление вплоть до максимальной аналитической частоты, представляющей собой только небольшую величину ниже половины частоты дискретизации и предпочтительно составляющей по меньшей мере одну четверть частоты дискретизации или типично выше.

Как проиллюстрировано, кодер работает без понижающей дискретизации, а декодер работает без повышающей дискретизации. Другими словами, аудиокодер в спектральной области выполнен с возможностью формировать спектральное представление, имеющее частоту Найквиста, заданную посредством частоты дискретизации первоначального входного аудиосигнала.

Кроме того, как проиллюстрировано на фиг. 3а, спектральный анализатор выполнен с возможностью анализировать спектральное представление начиная с начальной частоты заполнения интервалов и завершая максимальной частотой, представленной посредством максимальной частоты, включенной в спектральное представление, при этом спектральная часть, идущая от минимальной частоты вплоть до начальной частоты заполнения интервалов, принадлежит первому набору спектральных частей, и при этом дополнительная спектральная часть, к примеру, 304, 305, 306, 307, имеющая значения частоты выше частоты заполнения интервалов, дополнительно включена в первый набор первых спектральных частей.

Как указано, аудиодекодер 112 в спектральной области имеет такую конфигурацию, в которой максимальная частота, представленная посредством спектрального значения в первом декодированном представлении, равна максимальной частоте, включенной во временное представление, имеющее частоту дискретизации, при которой спектральное значение для максимальной частоты в первом наборе первых спектральных частей является нулем или отличается от нуля. В любом случае, для этой максимальной частоты в первом наборе спектральных компонентов существует коэффициент масштабирования для полосы частот коэффициентов масштабирования, который формируется и передается независимо от того, задаются или нет все спектральные значения в этой полосе частот коэффициентов масштабирования равными нулю, как пояснено в контексте фиг. 3а и 3b.

Следовательно, изобретение является преимущественным в том, что относительно других параметрических технологий для того, чтобы повышать эффективность сжатия, например, замещения шума и заполнения шумом (эти технологии служат исключительно для эффективного представления шумоподобного локального содержимого сигнала), изобретение обеспечивает возможность точного воспроизведения частоты тональных компонентов. К настоящему времени ни одна технология предшествующего уровня

техники не разрешает эффективное параметрическое представление произвольного содержимого сигнала посредством заполнения интервалов в спектре без ограничения фиксированного априорного разделения в полосе низких частот (LF) и полосе высоких частот (HF).

Варианты осуществления изобретаемой системы совершенствуют подходы предшествующего уровня техники и за счет этого предоставляют высокую эффективность сжатия, отсутствие либо только небольшое перцепционное раздражение и полную полосу пропускания аудиосигнала даже для низких скоростей передачи битов.

Общая система состоит из:

- полнополосного базового кодирования,
- интеллектуального заполнения интервалов (заполнения фрагментами или заполнения шумом),
- разреженных тональных частей в базовом кодере, выбранных посредством тональной маски,
- кодирования на основе объединенных стереопар для полной полосы частот, включающего в себя заполнение фрагментами,
- TNS для фрагмента,
- спектрального отбеливания в IGF-диапазоне.

Первый шаг к более эффективной системе заключается в том, чтобы устранять необходимость преобразования спектральных данных во вторую область преобразования, отличающуюся от области преобразования базового кодера. Поскольку большинство аудиокодеков, к примеру, AAC, например, используют MDCT в качестве базового преобразования, полезно также выполнять BWE в MDCT-области. Второе требование для BWE-системы заключается в необходимости сохранять тональную сетку, в силу чего сохраняются даже тональные HF-компоненты, и качество кодированного аудио за счет этого превосходит существующие системы. Чтобы обеспечивать выполнение обоих вышеуказанных требований для BWE-схемы, предложена новая система, называемая "интеллектуальным заполнением интервалов (IGF)". Фиг. 2b показывает блок-схему предложенной системы на стороне кодера, а фиг. 2a показывает систему на стороне декодера.

Фиг. 6a иллюстрирует устройство для декодирования кодированного аудиосигнала в другой реализации настоящего изобретения. Устройство для декодирования содержит аудиодекодер 602 в спектральной области для формирования первого декодированного представления первого набора спектральных частей и в качестве модуля 604 повторного формирования частоты, соединенного ниже аудиодекодера 602 в спектральной области для формирования восстановленной второй спектральной части с использованием первой спектральной части из первого набора первых спектральных частей. Как проиллюстрировано на 603, спектральные значения в первой спектральной части и во второй спектральной части являются остаточными спектральными прогнозными значениями. Чтобы преобразовывать эти остаточные спектральные прогнозные значения в полное спектральное представление, предоставляется спектральный прогнозный фильтр 606. Этот обратный прогнозный фильтр выполнен с возможностью осуществления обратного прогнозирования по частоте с использованием остаточных спектральных значений для первого набора первой частоты и восстановленных вторых спектральных частей. Спектральный обратный прогнозный фильтр 606 сконфигурирован посредством информации фильтра, включенной в кодированный аудиосигнал. Фиг. 6b иллюстрирует более подробную реализацию варианта осуществления по фиг. 6a. Остаточные спектральные прогнозные значения 603 вводятся в модуль 612

формирования частотных фрагментов, формирующий необработанные спектральные значения для полосы частот восстановления или для определенной второй частотной части, и эти необработанные данные, в данный момент имеющие разрешение, идентичное разрешению первого спектрального представления высокого разрешения, вводятся в спектральный формирователь 614. После этого спектральный формирователь придает определенную форму спектру с использованием информации огибающей, передаваемой в потоке битов, и данные определенной спектральной формы затем применяются к спектральному прогнозному фильтру 616, в завершение, формирующему кадр из полных спектральных значений с использованием информации 607 фильтра, передаваемой из кодера в декодер через поток битов.

На фиг. 6b, предполагается, что на стороне кодера, вычисление информации фильтра, передаваемой через поток битов и используемой через линию 607, выполняется после вычисления информации огибающей. Следовательно, другими словами, кодер, совпадающий с декодером по фиг. 6b, сначала должен вычислять остаточные спектральные значения, а затем должен вычислять информацию огибающей с помощью остаточных спектральных значений, как, например, проиллюстрировано на фиг. 7a. Тем не менее, другая реализация также является полезной для определенных реализаций, в которых информация огибающей вычисляется до выполнения TNS- или TTS-фильтрации на стороне кодера. В таком случае спектральный прогнозный фильтр 622 применяется до выполнения формирования спектра в блоке 624. Таким образом, другими словами, (полные) спектральные значения формируются до того, как применяется операция 624 формирования спектра.

Предпочтительно, вычисляется комплекснозначный TNS-фильтр или TTS-фильтр. Это проиллюстрировано на фиг. 7a. Исходный аудиосигнал вводится в комплексный MDCT-блок 702. Затем вычисление TTS-фильтра и TTS-фильтрация выполняются в комплексной области. После этого в блоке 706, вычисляется вспомогательная IGF-информация, а также вычисляется любая другая операция, к примеру, спектральный анализ для кодирования и т.д. Далее первый набор первой спектральной части, сформированной посредством блока 706, кодируется с помощью кодера на основе психоакустической модели, проиллюстрированного на 708, чтобы получать первый набор первых спектральных частей, указываемых в $X(k)$ на фиг. 7a, и все эти данные перенаправляются в мультиплексор 710 потоков битов.

На стороне декодера, кодированные данные вводятся в демультиплексор 720, чтобы разделять вспомогательную IGF-информацию, с одной стороны, вспомогательную TTS-информацию, с другой стороны, и кодированное представление первого набора первых спектральных частей.

Затем блок 724 используется для вычисления комплексного спектра из одного или более действительностнозначных спектров. Далее действительностнозначные и комплексные спектры вводятся в блок 726 для того, чтобы формировать восстановленные значения частоты во втором наборе вторых спектральных частей для полосы частот восстановления. После этого для полностью полученного и заполненного фрагментами полнополосного кадра выполняется обратная TTS-операция 728, и на стороне декодера выполняется конечная операция обратного комплексного MDCT в блоке 730. Таким образом, использование информации комплексного TNS-фильтра позволяет, не только при применении в полосе базовых частот или в отдельных полосах частот фрагментов, но и при применении на границах между полосой базовых частот и полосой частот фрагментов или на границах между полосами частот фрагментов, автоматически формировать обработку границ фрагментов, которая, в конечном счете, повторно

вводит спектральную корреляцию между фрагментами. Эта спектральная корреляция на границах фрагментов не получается только посредством формирования частотных фрагментов и выполнения регулирования спектральной огибающей для этих необработанных данных частотных фрагментов.

5 Фиг. 7с иллюстрирует сравнение исходного сигнала (левая панель) и расширенного сигнала без TTS. Можно видеть, что существуют сильные артефакты, проиллюстрированные посредством расширенных частей в диапазоне верхних частот, проиллюстрированном на 750. Тем не менее, это не возникает на фиг. 7е, когда идентичная спектральная часть на 750 сравнивается со связанным с артефактами компонентом 750 по фиг. 7с.

10 Варианты осуществления или изобретаемая система кодирования аудио используют основную долю доступной скорости передачи битов для того, чтобы кодировать на основе формы сигналов только перцепционно самую релевантную структуру сигнала в кодере, и результирующие интервалы отсутствия сигнала в спектре заполняются в декодере содержимого сигнала, который примерно аппроксимирует исходный спектр. 15 Очень ограниченный битовый бюджет расходуется для того, чтобы управлять так называемым интеллектуальным заполнением интервалов (IGF) в спектре на основе параметров посредством выделенной вспомогательной информации, передаваемой из кодера в декодер.

20 Хранение или передача аудиосигналов зачастую подчиняются строгим ограничениям скорости передачи битов. В прошлом, кодеры принудительно существенно уменьшали полосу пропускания передаваемого аудиосигнала, когда была доступна только очень низкая скорость передачи битов. Современные аудиокодеки в наше время могут кодировать широкополосные сигналы посредством использования таких способов 25 расширения полосы пропускания (BWE) как репликация полосы пропускания спектра (SBR) [1]. Эти алгоритмы основываются на параметрическом представлении высокочастотного содержимого (HF), который формируется из кодированной на основе формы сигналов низкочастотной части (LF) декодированного сигнала, посредством транспозиции в спектральную HF-область ("наложения") и применения постобработки 30 на основе параметров. В BWE-схемах, восстановление спектральной HF-области выше данной так называемой частоты разделения зачастую основано на спектральном наложении. Типично, HF-область состоит из нескольких смежных наложений, и каждое из этих наложений получается из полосовых (BP) областей LF-спектра ниже данной частоты разделения. Системы предшествующего уровня техники эффективно выполняют 35 наложение в представлении на основе гребенки фильтров посредством копирования набора смежных подполосных коэффициентов из исходной в целевую область.

Если BWE-система реализуется в области гребенки фильтров или частотно-временного преобразования, предусмотрена только ограниченная возможность 40 управлять временной формой сигнала расширения полосы пропускания. Типично, степень временной детализации ограничена посредством размера перескока, используемого между смежными окнами преобразования на основе кодирования со взвешиванием. Это может приводить к нежелательным опережающим или запаздывающим эхо в спектральном BWE-диапазоне.

Из перцепционного кодирования аудио известно, что форма временной огибающей 45 аудиосигнала может быть восстановлена посредством использования спектральных технологий фильтрации, таких как временное формирование огибающей (TNS)[14]. Тем не менее, TNS-фильтр, известный из предшествующего уровня техники, представляет собой действительнoзначный фильтр в действительнoзначных спектрах. Такой

действительнозначный фильтр в действительнозначных спектрах может серьезно искажаться посредством артефактов наложения спектров, в частности, если базовое действительное преобразование представляет собой модифицированное дискретное косинусное преобразование (MDCT).

Формирование фрагментов временной огибающей применяет комплексную фильтрацию к комплекснозначным спектрам, к примеру, полученным, например, из комплексного модифицированного дискретного косинусного преобразования (CMDCT). В силу этого, не допускаются артефакты наложения спектров.

Временное формирование фрагментов состоит из:

- оценки коэффициентов комплексной фильтрации и применения сглаживающего фильтра к спектру исходного сигнала в кодере,
- передачи коэффициентов фильтрации во вспомогательной информации,
- применения формирующего фильтра в заполненном фрагментами восстановленном спектре в декодере.

Изобретение дополняет технологию предшествующего уровня техники, известную из кодирования аудио с преобразованием, в частности, временного формирования шума (TNS), посредством линейного прогнозирования вдоль направления частоты, для использования модифицированным способом в контексте расширения полосы пропускания.

Дополнительно, изобретаемый алгоритм расширения полосы пропускания основан на интеллектуальном заполнении интервалов (IGF), но использует избыточно дискретизированное комплекснозначное преобразование (CMDCT), в противоположность стандартной IGF-конфигурации, которая основывается на действительнозначном критически дискретизированном MDCT-представлении сигнала.

CMDCT может рассматриваться в качестве комбинации MDCT-коэффициентов в действительной части и MDST-коэффициентов в мнимой части каждого комплекснозначного спектрального коэффициента.

Хотя новый подход описывается в контексте IGF, изобретаемая обработка может использоваться в сочетании с любым BWE-способом, который основан на представлении на основе гребенки фильтров аудиосигнала.

В этом новом контексте линейное прогнозирование вдоль направления частоты используется не в качестве временного формирования шума, а вместо этого в качестве технологии временного формирования фрагментов (TTS). Переименование оправдано тем фактом, что заполненным фрагментами компонентам сигнала придается определенная временная форма посредством TTS, в противоположность формированию шума квантования посредством TNS в перцепционных кодеках с преобразованием предшествующего уровня техники.

Фиг. 7а показывает блок-схему BWE-кодера с использованием IGF- и нового TTS-подхода.

Таким образом, базовая схема кодирования работает следующим образом:

- вычисление CMDCT сигнала $x(n)$ временной области для того, чтобы получить сигнал $X(k)$ частотной области,
- вычисление комплекснозначного TTS-фильтра,
- получение вспомогательной информации для BWE и удаление спектральной информации, которая должна реплицироваться посредством декодера,
- применение квантования с использованием психоакустического модуля (PAM),
- сохранение/передача данных, передаются только действительнозначные MDCT-коэффициенты.

Фиг. 7b показывает соответствующий декодер. Он, главным образом, выполняет в обратном порядке этапы, осуществляемые в кодере.

Здесь, базовая схема декодирования работает следующим образом:

- оценка MDST-коэффициентов из MDCT-значений (эта обработка добавляет одну задержку блочного декодера) и комбинирование MDCT- и MDST-коэффициентов в комплекснозначные CMDCT-коэффициенты,

- выполнение заполнения мозаичными фрагментами с их постобработкой,

- применение обратной TTS-фильтрации с помощью передаваемых коэффициентов TTS-фильтрации,

- вычисление обратного CMDCT.

Следует отметить, что, альтернативно, порядок TTS-синтеза и IGF-постобработки также может быть изменен на противоположный в декодере, если соответствующим образом изменяются на противоположное TTS-анализ и оценка IGF-параметров в кодере.

Для эффективного кодирования с преобразованием предпочтительно так называемые "длинные блоки" приблизительно 20 мс должны использоваться для того, чтобы достигать обоснованного усиления преобразования. Если сигнал в таком длинном блоке содержит переходные части, слышимые опережающие и запаздывающие эхо возникают в восстановленных полосах спектра вследствие заполнения мозаичными фрагментами. Фиг. 7c показывает типичные эффекты опережающего и запаздывающего эхо, которые искажают переходные части вследствие IGF. На левой панели по фиг. 7c, показана спектрограмма исходного сигнала, а на правой панели показана спектрограмма заполненной мозаичными фрагментами сигнала без изобретаемой TTS-фильтрации. В этом примере, начальная IGF-частота $f_{IGFstart}$ или f_{split} между полосой базовых частот и заполненной мозаичными фрагментами полосой частот выбрана в качестве $f_s/4$. В правой панели по фиг. 7c, различные опережающие и запаздывающие эхо являются видимым с окружением переходных частей, в частности, заметных в верхнем спектральном конце реплицируемой частотной области.

Основная задача модуля TTS состоит в том, чтобы ограничивать эти нежелательные компоненты сигнала в окрестности вокруг переходной части и за счет этого скрывать их во временной области, управляемой посредством эффекта временного маскирования человеческого восприятия. Следовательно, требуемые коэффициенты TTS-прогнозирования вычисляются и применяются с использованием "прямого прогнозирования" в CMDCT-области.

В варианте осуществления, который комбинирует TTS и IGF в кодек, важно совмещать определенные TTS-параметры и IGF-параметры таким образом, что мозаичный IGF-фрагмент либо полностью фильтруется посредством одного TTS-фильтра (сглаживающего или формирующего фильтра), либо нет. Следовательно, все частоты TTSstart[...] или TTSstop[...] не должны содержаться в мозаичном IGF-фрагменте, а вместо этого должны совмещаться с соответствующими частотами $f_{IGF...}$. Фиг. 7d показывает пример рабочих TTS- и IGF-областей для набора из трех TTS-фильтров.

Конечная TTS-частота регулируется до конечной частоты инструментального IGF-средства, которая выше $f_{IGFstart}$. Если TTS использует более одного фильтра, необходимо обеспечивать то, что частота разделения между двумя TTS-фильтрами должна совпадать с частотой разбиения IGF. В противном случае, один субфильтр TTS должен проходить по $f_{IGFstart}$, что приводит к нежелательным артефактам, таким как избыточное

формирование.

В варианте реализации, проиллюстрированном на фиг. 7а и фиг. 7b, следует принимать дополнительные меры для того, чтобы в этом декодере корректно регулировать IGF-энергии. Это в особенности имеет место, если в ходе TTS и обработки IGF, различные
 5 TTS-фильтры, имеющие различные усиления для прогнозирования, применяются к исходной области (в качестве сглаживающего фильтра) и к целевой спектральной области (в качестве формирующего фильтра, который не является точным аналогом упомянутого сглаживающего фильтра) одного мозаичного IGF-фрагмента. В этом случае, коэффициент усиления для прогнозирования двух применяемых TTS-фильтров
 10 более не равен единице, и следовательно, должно применяться энергетическое регулирование посредством этого коэффициента.

В альтернативном варианте реализации, порядок IGF-постобработки и TTS изменяется на противоположный. В декодере это означает то, что энергетическое регулирование посредством IGF-постобработки вычисляется после TTS-фильтрации и в силу этого
 15 представляет собой конечный этап обработки перед синтезирующим преобразованием. Следовательно, независимо от применения различных усилений TTS-фильтра к одному мозаичному фрагменту в ходе кодирования, конечная энергия всегда регулируется корректно посредством обработки IGF.

На стороне декодера, коэффициенты TTS-фильтрации снова применяются к полному
 20 спектру, т.е. к базовому спектру, расширенному посредством повторно сформированного спектра. Применение TTS необходимо для того, чтобы снова формировать временную огибающую повторно сформированного спектра таким образом, что она совпадает с огибающей исходного сигнала. Таким образом, уменьшаются показанные опережающие эхо. Помимо этого, оно по-прежнему временно
 25 придет определенную форму шуму квантования в сигнале ниже $f_{IGFstart}$, что является стандартным для унаследованного TNS.

В унаследованных кодерах, спектральное наложение на аудиосигнал (например, SBR) нарушает спектральную корреляцию на границах наложений и в силу этого искажает временную огибающую аудиосигнала посредством введения дисперсии.
 30 Следовательно, другое преимущество выполнения заполнения мозаичными IGF-фрагментами для остаточного сигнала заключается в том, что после применения формирующего TTS-фильтра границы мозаичных фрагментов прозрачно коррелированы, приводя к более достоверному временному воспроизведению сигнала.

Результат соответствующим образом обработанного сигнала показан на фиг. 7е. В
 35 сравнении нефильтрованной версии (фиг. 7с, правая панель) TTS-фильтрованный сигнал показывает хорошее уменьшение нежелательных опережающих и запаздывающих эхо (фиг. 7е, правая панель).

Кроме того, как пояснено, фиг. 7а иллюстрирует кодер, совпадающий с декодером по фиг. 7b или с декодером по фиг. 6а. По существу, устройство для кодирования
 40 аудиосигнала содержит временно-спектральный преобразователь, к примеру, 702 для преобразования аудиосигнала в спектральное представление. Спектральное представление может представлять собой спектральное представление действительных значений или, как проиллюстрировано на этапе 702, спектральное представление комплексных значений. Кроме того, предоставляется прогнозный фильтр, к примеру,
 45 704 для выполнения прогнозирования по частоте, чтобы формировать остаточные спектральные значения, при этом прогнозный фильтр 704 задается посредством информации прогнозного фильтра, извлекаемой из аудиосигнала и перенаправленной в мультиплексор 710 потоков битов, как проиллюстрировано на 714 на фиг. 7а. Кроме

того, предоставляется аудиокодер, такой как психоакустически регулируемый аудиокодер 704. Аудиокодер выполнен с возможностью кодирования первого набора первых спектральных частей остаточных спектральных значений для того, чтобы получать кодированный первый набор первых спектральных значений. Кроме того, параметрический кодер, такой как кодер, проиллюстрированный на 706 на фиг. 7а, предоставляется для кодирования второго набора вторых спектральных частей. Предпочтительно, первый набор первых спектральных частей кодируется с более высоким спектральным разрешением по сравнению со вторым набором вторых спектральных частей.

В завершение, как проиллюстрировано на фиг. 7а, предоставляется интерфейс вывода для вывода кодированного сигнала, содержащего параметрически кодированный второй набор вторых спектральных частей, кодированный первый набор первых спектральных частей и информацию фильтра, проиллюстрированную в качестве "вспомогательной TTS-информации" на 714 на фиг. 7а.

Предпочтительно, прогнозный фильтр 704 содержит модуль вычисления информации фильтра, выполненный с возможностью использования спектральных значений спектрального представления для вычисления информации фильтра. Кроме того, прогнозный фильтр выполнен с возможностью вычисления остаточных спектральных значений с использованием идентичных спектральных значений спектрального представления, используемого для вычисления информации фильтра.

Предпочтительно, TTS-фильтр 704 сконфигурирован аналогичным образом, как известно для аудиокодеров предшествующего уровня техники, применяющих инструментальное TNS-средство в соответствии с AAC-стандартом.

Ниже поясняется дополнительная реализация с использованием двухканального декодирования в контексте фиг. 8а-8е. Кроме того, следует обратиться к описанию соответствующих элементов в контексте фиг. 2а, 2b (объединенное канальное кодирование 228 и объединенное канальное декодирование 204).

Фиг. 8а иллюстрирует аудиодекодер для формирования декодированного двухканального сигнала. Аудиодекодер содержит четыре аудиодекодера 802 для декодирования кодированного двухканального сигнала, чтобы получать первый набор первых спектральных частей, и дополнительно параметрический декодер 804 для предоставления параметрических данных для второго набора вторых спектральных частей, и дополнительно, двухканальный идентификатор, идентифицирующий или первое или второе отличное двухканальное представление для вторых спектральных частей. Кроме того, модуль 806 повторного формирования частоты предоставляется для повторного формирования второй спектральной части в зависимости от первой спектральной части из первого набора первых спектральных частей и параметрических данных для второй части и двухканального идентификатора для второй части. Фиг. 8b иллюстрирует различные комбинации для двухканальных представлений в исходном диапазоне и целевом диапазоне. Исходный диапазон может находиться в первом двухканальном представлении, и целевой диапазон также может находиться в первом двухканальном представлении. Альтернативно, исходный диапазон может находиться в первом двухканальном представлении, а целевой диапазон может находиться во втором двухканальном представлении. Кроме того, исходный диапазон может находиться во втором двухканальном представлении, а целевой диапазон может находиться в первом двухканальном представлении, как указано в третьем столбце по фиг. 8b. В завершение, исходный диапазон и целевой диапазон могут находиться во втором двухканальном представлении. В варианте осуществления, первое двухканальное

представление представляет собой отдельное двухканальное представление, в котором два канала двухканального сигнала представляются отдельно. В таком случае, второе двухканальное представление представляет собой объединенное представление, в котором два канала двухканального представления представляются совместно, т.е. в котором требуется последующая обработка или преобразование представления для того, чтобы повторно вычислять отдельное двухканальное представление по мере необходимости для вывода в соответствующие динамики.

В реализации, первое двухканальное представление может представлять собой левое/правое (L/R) представление, а второе двухканальное представление представляет собой объединенное стереопредставление. Тем не менее, другие двухканальные представления, отличные от левого/правого или M/S, или стереопрогнозирования, могут применяться и использоваться для настоящего изобретения.

Фиг. 8с иллюстрирует блок-схему последовательности операций способа для операций, выполняемых посредством аудиодекодера по фиг. 8а. На этапе 812 аудиодекодер 802 выполняет декодирование исходного диапазона. Исходный диапазон может содержать, относительно фиг. 3а, полосы SCB1-SCB3 частот коэффициентов масштабирования. Кроме того, может быть двухканальный идентификатор для каждой полосы частот коэффициентов масштабирования, и полоса 1 частот коэффициентов масштабирования, например, может находиться в первом представлении (к примеру, L/R), а третья полоса частот коэффициентов масштабирования может находиться во втором двухканальном представлении, таком как M/S или низведение/остаток прогнозирования. Таким образом, этап 812 может приводить к различным представлениям для различных полос частот. Затем на этапе 814, модуль 806 повторного формирования частоты выполнен с возможностью выбора исходного диапазона для повторного формирования частоты. На этапе 816, модуль 806 повторного формирования частоты затем проверяет представление исходного диапазона, и в блоке 818 модуль 806 повторного формирования частоты сравнивает двухканальное представление исходного диапазона с двухканальным представлением целевого диапазона. Если оба представления являются идентичными, модуль 806 повторного формирования частоты предоставляет отдельное повторное формирование частоты для каждого канала двухканального сигнала. Тем не менее, когда оба представления, обнаруженные в блоке 818, не являются идентичными, в таком случае извлекается последовательность 824 сигналов, и блок 822 вычисляет другое двухканальное представление из исходного диапазона и использует это вычисленное другое двухканальное представление для повторного формирования целевого диапазона. Таким образом, декодер по фиг. 8а позволяет повторно формировать целевой диапазон, указываемый как имеющий второй двухканальный идентификатор, с использованием исходного диапазона, находящегося в первом двухканальном представлении. Естественно, настоящее изобретение дополнительно дает возможность повторно формировать целевой диапазон с использованием исходного диапазона, имеющего идентичный двухканальный идентификатор. Кроме того, дополнительно настоящее изобретение дает возможность повторно формировать целевой диапазон, имеющий двухканальный идентификатор, указывающий объединенное двухканальное представление, и затем преобразовывать это представление в отдельное канальное представление, требуемое для хранения или передачи в соответствующие громкоговорители для двухканального сигнала.

Следует подчеркнуть, что два канала двухканального представления могут представлять собой два стереоканала, к примеру, левый канал и правый канал. Тем не менее, сигнал также может представлять собой многоканальный сигнал, имеющий,

например, пять каналов и сабвуферный канал, или имеющий еще большее число каналов. После этого, может выполняться попарная двухканальная обработка, как пояснено в контексте фиг. 8а-8е, при этом пары, например, могут представлять собой левый канал и правый канал, левый канал объемного звучания и правый канал объемного звучания и центральный канал и LFE-(субвуферный) канал. Любые другие спаривания могут использоваться для того, чтобы представлять, например, шесть входных каналов посредством трех процедур двухканальной обработки.

Фиг. 8d иллюстрирует блок-схему изобретаемого декодера, соответствующего фиг. 8а. Декодер 830 исходного диапазона (или базовый) может соответствовать аудиодекодеру 802. Другие блоки 832, 834, 836, 838, 840, 842 и 846 могут быть частями модуля 806 повторного формирования частоты по фиг. 8а. В частности, блок 832 представляет собой модуль преобразования представлений для преобразования представлений исходного диапазона в отдельных полосах частот, так что на выходе блока 832 присутствует полный набор исходного диапазона в первом представлении, с одной стороны, и во втором двухканальном представлении, с другой стороны. Эти два полных представления исходного диапазона могут сохраняться в устройстве 834 хранения данных для обоих представлений исходного диапазона.

Затем блок 836 применяет формирование частотных мозаичных фрагментов с использованием, в качестве ввода, идентификатора исходного диапазона и дополнительно с использованием, в качестве ввода, двухканального идентификатора для целевого диапазона. На основе двухканального идентификатора для целевого диапазона модуль формирования частотных мозаичных фрагментов осуществляет доступ к устройству 834 хранения данных и принимает двухканальное представление исходного диапазона, совпадающее с двухканальным идентификатором для целевого диапазона, вводимого в модуль формирования частотных мозаичных фрагментов на 835. Таким образом, когда двухканальный идентификатор для целевого диапазона указывает объединенную стереообработку, в таком случае модуль 836 формирования частотных мозаичных фрагментов осуществляет доступ к устройству 834 хранения данных для того, чтобы получать объединенное стереопредставление исходного диапазона, указываемого посредством идентификатора 833 исходного диапазона.

Модуль 836 формирования частотных мозаичных фрагментов выполняет эту операцию для каждого целевого диапазона, и вывод модуля формирования частотных мозаичных фрагментов является таким, что присутствует каждый канал канального представления, идентифицированный посредством двухканального идентификатора. Затем выполняется регулирование огибающей посредством модуля 838 регулирования огибающей. Регулирование огибающей выполняется в двухканальной области, идентифицированной посредством двухканального идентификатора. С этой целью, требуются параметры регулирования огибающей, и эти параметры передаются из кодера в декодер в идентичном двухканальном представлении, как описано выше. Когда двухканальный идентификатор в целевом диапазоне, который должен быть обработан посредством модуля регулирования огибающей, имеет двухканальный идентификатор, указывающий другое двухканальное представление по сравнению с данными огибающей для этого целевого диапазона, в таком случае модуль 840 преобразования параметров преобразует параметры огибающей в требуемое двухканальное представление. Когда, например, двухканальный идентификатор для одной полосы частот указывает объединенное стереокодирование, и когда параметры для этого целевого диапазона переданы в качестве параметров L/R-огибающей, в таком случае модуль преобразования параметров вычисляет параметры объединенной

стереоогibaющей из параметров L/R-огibaющей так, как описано, так что корректное параметрическое представление используется для регулирования спектральной огibaющей целевого диапазона.

В другом предпочтительном варианте осуществления, параметры огibaющей уже переданы в качестве параметров объединенного стерео, когда объединенное стерео используется в целевой полосе частот.

Когда предполагается, что ввод в модуль 838 регулирования огibaющей представляет собой набор целевых диапазонов, имеющих различные двухканальные представления, в таком случае вывод модуля 838 регулирования огibaющей также представляет собой набор целевых диапазонов в различных двухканальных представлениях. Когда, целевой диапазон имеет объединенное представление, к примеру, M/S, в таком случае этот целевой диапазон обрабатывается посредством модуля 842 преобразования представлений для вычисления отдельного представления, требуемого для хранения или передачи в громкоговорители. Тем не менее, когда целевой диапазон уже имеет отдельное представление, извлекается последовательность 844 сигналов, и модуль 842 преобразования представлений обходится. На выходе блока 842 получается двухканальное спектральное представление, представляющее собой отдельное двухканальное представление, которое затем может дополнительно обрабатываться, как указано посредством блока 846, причем эта последующая обработка, например, может представлять собой частотно-временное преобразование или любую другую требуемую обработку.

Предпочтительно, вторые спектральные части соответствуют полосам частот, и двухканальный идентификатор предоставляется в качестве матрицы флагов, соответствующих таблице по фиг. 8b, в которой предусмотрен один флаг для каждой полосы частот. После этого параметрический декодер выполнен с возможностью проверять то, задается флаг или нет, и управлять модулем 106 повторного формирования частоты в соответствии с флагом таким образом, чтобы использовать либо первое представление, либо второе представление первой спектральной части.

В варианте осуществления, только диапазон восстановления начиная с начальной IGF-частоты 309 по фиг. 3a имеет двухканальные идентификаторы для различных полос частот восстановления. В дополнительном варианте осуществления, это также применимо для частотного диапазона ниже начальной IGF-частоты 309.

В дополнительном варианте осуществления, идентификатор исходной полосы частот и идентификатор целевой полосы частот могут быть адаптивно определены посредством анализа подобию. Тем не менее, изобретаемая двухканальная обработка также может применяться, когда существует фиксированное ассоциирование исходного диапазона с целевым диапазоном. Исходный диапазон может использоваться для повторного создания, относительно частоты, более широкого целевого диапазона либо посредством операции заполнения частотными мозаичными фрагментами гармоник, либо посредством операции заполнения частотными мозаичными фрагментами перезаписи с использованием двух или более операций заполнения частотными мозаичными фрагментами, аналогично обработке для нескольких наложений, известной из высокоэффективной AAC-обработки.

Фиг. 8e иллюстрирует аудиокодер для кодирования двухканального аудиосигнала. Кодер содержит временно-спектральный преобразователь 860 для преобразования двухканального аудиосигнала в спектральное представление. Кроме того, спектральный анализатор 866 для преобразования аудиосигнала с двухканальным аудиоканалом в спектральное представление. Кроме того, спектральный анализатор 866 предоставляется

для выполнения анализа, чтобы определять то, какие спектральные части должны быть кодированы с высоким разрешением, т.е. выявлять первый набор первых спектральных частей и дополнительно выявлять второй набор вторых спектральных частей.

Кроме того, двухканальный анализатор 864 предоставляется для анализа второго набора вторых спектральных частей для того, чтобы определять двухканальный идентификатор, идентифицирующий либо первое двухканальное представление, либо второе двухканальное представление.

В зависимости от результата двухканального анализатора полоса частот во втором спектральном представлении параметризуется с использованием первого двухканального представления или второго двухканального представления, и это выполняется посредством кодера 868 параметров. Базовый частотный диапазон, т.е. полоса частот ниже начальной IGF-частоты 309 по фиг. 3а, кодируется посредством базового кодера 870. Результат блоков 868 и 870 вводится в интерфейс 872 вывода. Как указано, двухканальный анализатор предоставляет двухканальный идентификатор либо для каждой полосы частот выше начальной IGF-частоты, либо для всего целого частотного диапазона, и этот двухканальный идентификатор также перенаправляется в интерфейс 872 вывода, так что эти данные также включаются в кодированный сигнал 873, выводимый посредством интерфейса 872 вывода.

Кроме того, предпочтительно, чтобы аудиокодер содержал модуль 862 преобразования для каждой полосы частот. На основе решения двухканального анализатора 862 выходной сигнал временно-спектрального преобразователя 862 преобразован в представление, указываемое посредством двухканального анализатора и, в частности, посредством двухканального идентификатора 835. Таким образом, вывод модуля 862 преобразования для каждой полосы частот представляет собой набор полос частот, при этом каждая полоса частот может иметь или первое двухканальное представление, или второе отличное двухканальное представление. Когда настоящее изобретение применяется в полной полосе частот, т.е. когда исходный диапазон и диапазон восстановления обрабатываются посредством модуля преобразования для каждой полосы частот, спектральный анализатор 860 может анализировать это представление. Тем не менее, альтернативно, спектральный анализатор 860 также может анализировать сигнал, выводимый посредством временно-спектрального преобразователя, как указано посредством линии 861 управления. Таким образом, спектральный анализатор 860 может применять предпочтительный анализ тональности либо к выводу модуля 862 преобразования для каждой полосы частот, либо к выводу временно-спектрального преобразователя 860 до обработки посредством модуля 862 преобразования для каждой полосы частот. Кроме того, спектральный анализатор может применять идентификацию наилучшего совпадающего исходного диапазона для определенного целевого диапазона либо к результату модуля 862 преобразования для каждой полосы частот, либо к результату временно-спектрального преобразователя 860.

Далее следует обратиться к фиг. 9а на 9d для иллюстрации предпочтительного вычисления значений информации энергии, уже поясненных в контексте фиг. 3а и фиг. 3b.

Современные аудиокодеры предшествующего уровня применяют различные технологии для того, чтобы минимизировать объем данных, представляющих данный аудиосигнал. Аудиокодеры, такие как USAC [1], применяют частотное-временное преобразование, такое как MDCT, для того чтобы получать спектральное представление данного аудиосигнала. Эти MDCT-коэффициенты квантуются с использованием

психоакустических аспектов слуховой системы человека. Если доступная скорость передачи битов снижается, квантование становится более приблизительным, вводя большие количества обнуленных спектральных значений, что приводит к слышимым артефактам на стороне декодера. Чтобы повышать перцепционное качество, современные декодеры заполняют эти обнуленные спектральные части случайным шумом. IGF-способ собирает мозаичные фрагменты из оставшегося ненулевого сигнала, чтобы заполнять эти интервалы отсутствия сигнала в спектре. Для перцепционного качества декодированного аудиосигнала крайне важно, чтобы сохранялись спектральная огибающая и распределение энергии спектральных коэффициентов. Способ энергетического регулирования, представленный здесь, использует передаваемую вспомогательную информацию для того, чтобы восстанавливать спектральную MDCT-огибающую аудиосигнала.

В eSBR [15] аудиосигнал понижающе дискретизируется, по меньшей мере, на коэффициент два, и высокочастотная часть спектра полностью обнуляется [1, 17]. Эта удаленная часть заменяется посредством параметрических технологий, eSBR, на стороне декодера. eSBR подразумевает использование дополнительного преобразования, QMF-преобразования, которое используется для того, чтобы заменять пустую высокочастотную часть и повторно дискретизировать аудиосигнал [17]. Это повышает как вычислительную сложность, так и потребление запоминающего устройства в аудиокодере.

USAC-кодер [15] предлагает возможность заполнять спектральные провалы (обнуленные спектральные линии) случайным шумом, но имеет следующие недочеты: случайный шум не может сохранять точную временную структуру переходного сигнала, и он не может сохранять гармоническую структуру тонального сигнала.

Область, в которой eSBR работает на стороне декодера, полностью удалена посредством кодера [1]. Следовательно, eSBR предрасположена к удалению тональных линий в высокочастотной области или искажению гармонических структур исходного сигнала. Поскольку частотное QMF-разрешение eSBR является очень низким, и повторная вставка синусоидальных компонентов является возможной только при приблизительном разрешении базовой гребенки фильтров, повторное формирование тональных компонентов в eSBR в реплицируемом частотном диапазоне имеет очень низкую точность.

eSBR использует технологии для того, чтобы регулировать энергии накладываемых областей, т.е. регулирование спектральной огибающей [1]. Эта технология использует передаваемые значения энергии в частотно-временной QMF-сетке для того, чтобы повторно придавать определенную форму спектральной огибающей. Эта технология предшествующего уровня техники не обрабатывает частично удаленные спектры, и вследствие высокого временного разрешения, она предрасположена к необходимости относительно большого числа битов для того, чтобы либо передавать надлежащие значения энергии, либо применять приблизительное квантование к значениям энергии.

Способу IGF не требуется дополнительное преобразование, поскольку он использует унаследованное MDCT-преобразование, которое вычисляется так, как описано в [15].

Способ энергетического регулирования, представленный здесь, использует вспомогательную информацию, сформированную посредством кодера, для того чтобы восстанавливать спектральную огибающую аудиосигнала. Эта вспомогательная информация формируется посредством кодера, как указано ниже:

а) Применение MDCT-преобразования на основе кодирования со взвешиванием к входному аудиосигналу [16, раздел 4.6], необязательное вычисление MDCT на основе

кодирования со взвешиванием или оценка MDCT на основе кодирования со взвешиванием из вычисленного MDCT.

б) Применение TNS/TTS к MDCT-коэффициентам [15, раздел 7.8].

5 с) Вычисление средней энергии для каждой полосы MDCT-частот коэффициентов масштабирования выше начальной IGF-частоты ($f_{IGFstart}$) вплоть до конечной IGF-частоты ($f_{IGFstop}$).

д) Квантование средних значений энергии.

$f_{IGFstart}$ и $f_{IGFstop}$ являются предоставленным пользователем параметрами.

10 Вычисленные значения из этапа с) и d) кодируются без потерь и передаются в качестве вспомогательной информации с потоком битов в декодер.

Декодер принимает передаваемые значения и использует их для того, чтобы регулировать спектральную огибающую.

а) Деквантование передаваемых MDCT-значений.

15 б) Применение унаследованного заполнения USAC-шумом, если передано в служебных сигналах.

с) Применение заполнения мозаичными IGF-фрагментами.

д) Деквантование передаваемых значений энергии.

20 е) Регулирование спектральной огибающей в расчете на каждую полосу частот коэффициентов масштабирования.

ф) Применение TNS/TTS, если передано в служебных сигналах.

Пусть $\hat{x} \in \mathbb{R}^N$ является MDCT-преобразованным действительнзначным спектральным представлением кодированного со взвешиванием аудиосигнала с длиной окна кодирования со взвешиванием в $2N$. Это преобразование описывается в [16]. Кодер
25 необязательно применяет TNS к \hat{x} .

В [16, 4.6.2] описывается сегмент \hat{x} в полосах частот коэффициентов масштабирования. Полосы частот коэффициентов масштабирования представляют собой набор набора индексов и обозначаются в этом тексте с помощью scb .

30 Пределы каждого scb_k with $k = 0, 1, 2, \dots, max_sfb$ задаются посредством матрицы swb_offset [16, 4.6.2] swb_offset [2, 4.6.2], где $swb_offset[k]$ и $swb_offset[k+1]-1$ задают первый и последний индекс для линии наименьших и наибольших спектральных коэффициентов, содержащейся в scb_k . Обозначим полосу частот коэффициентов масштабирования:

35 $scb_k := \{swb_offset[k], 1+swb_offset[k], 2+swb_offset[k], \dots, swb_offset[k+1]-1\}$

Если инструментальное IGF-средство используется посредством кодера, пользователь задает начальную IGF-частоту и конечную IGF-частоту. Эти два значения преобразуются в индекс $igfStartSfb$ и $igfStopSfb$. полосы частот коэффициентов масштабирования метода с наилучшим приближением. Оба передаются в служебных сигналах в потоке
40 битов в декодер.

[16] описывает преобразование как для длинных блоков, так и для коротких блоков. Для длинных блоков только один набор спектральных коэффициентов вместе с одним набором коэффициентов масштабирования передается в декодер. Для коротких блоков вычисляются восемь коротких окон кодирования со взвешиванием с восемью
45 различными наборами спектральных коэффициентов. Чтобы сокращать скорость передачи битов, коэффициенты масштабирования этих восьми окон кодирования со взвешиванием коротких блоков группируются посредством кодера.

В случае IGF способ, представленный здесь, использует унаследованные полосы

частот коэффициентов масштабирования для того, чтобы группировать спектральные значения, которые передаются в декодер:

$$E_k = \sqrt{\frac{1}{|scb_k|} \sum_{i \in scb_k} \hat{x}_i^2}$$

где $k = igfStartSfb, 1 + igfStartSfb, 2 + igfStartSfb, \dots, igfEndSfb$.

Для квантования,

$$\tilde{E}_k = nINT(4 \log_2(E_k))$$

вычисляется. Все значения \tilde{E}_k передаются в декодер.

Предположим, что кодер решает группировать num_window_group наборов коэффициентов масштабирования.

Обозначим как w этот сегмент группировки набора $\{0, 1, 2, \dots, 7\}$, который представляет собой индексы восьми коротких окон кодирования со взвешиванием; w_l обозначает l -ый поднабор w , где l обозначает индекс группы окон кодирования со взвешиванием, $0 \leq l < num_window_group$.

Для вычисления коротких блоков определяемая пользователем начальная/конечная IGF-частота преобразуется в надлежащие полосы частот коэффициентов масштабирования. Тем не менее, для простоты также обозначается для коротких блоков $k = igfStartSfb, 1 + igfStartSfb, 2 + igfStartSfb, \dots, igfEndSfb$.

Вычисление IGF-энергии использует информацию группировки для того, чтобы группировать значения $E_{k,l}$:

$$E_{k,l} := \sqrt{\frac{1}{|w_l|} \sum_{j \in w_l} \frac{1}{|scb_k|} \sum_{i \in scb_k} \hat{x}_{j,i}^2}$$

Для квантования,

$$\tilde{E}_{k,l} = nINT(4 \log_2(E_{k,l}))$$

вычисляется. Все значения $\tilde{E}_{k,l}$ передаются в декодер.

Вышеуказанные формулы кодирования работают с использованием только действительнзначных MDCT-коэффициентов \hat{x} . Для того, чтобы получать более стабильное распределение энергии в IGF-диапазоне, т.е. чтобы уменьшать временные амплитудные флуктуации, альтернативный способ может использоваться для того, чтобы вычислять значения \tilde{E}_k :

Пусть $\hat{x}_r \in \mathbb{R}^N$ является MDCT-преобразованным действительнзначным спектральным представлением кодированного со взвешиванием аудиосигнала с длиной окна кодирования со взвешиванием $2N$, а $\hat{x}_i \in \mathbb{R}^N$ является действительнзначным MDST-преобразованным спектральным представлением идентичной части аудиосигнала. Спектральное MDST-представление \hat{x}_i может быть либо вычислено точно, либо оценено из \hat{x}_r . $\hat{c} := (\hat{x}_r, \hat{x}_i) \in \mathbb{C}^N$ обозначает комплексное спектральное представление кодированного со взвешиванием аудиосигнала, имеющее \hat{x}_r в качестве действительной части и \hat{x}_i в качестве мнимой части. Кодер необязательно применяет TNS к \hat{x}_r и \hat{x}_i .

Далее, энергия исходного сигнала в IGF-диапазоне может измеряться с помощью:

$$E_{ok} = \frac{1}{|scb_k|} \sum_{i \in scb_k} \hat{c}_i^2$$

Действительно- и комплекснозначные энергии полосы частот восстановления, т.е. мозаичного фрагмента, который должен использоваться на стороне декодера при восстановлении IGF-диапазона scb_k , вычисляются следующим образом:

$$E_{tk} = \frac{1}{|scb_k|} \sum_{i \in tr_k} \hat{c}_i^2, \quad E_{rk} = \frac{1}{|scb_k|} \sum_{i \in tr_k} \hat{x}_{ri}^2,$$

где tr_k представляет собой набор индексов – ассоциированный диапазон исходных мозаичных фрагментов в зависимости от scb_k . В двух вышеприведенных формулах, вместо набора индексов scb_k , может использоваться набор $\overline{scb_k}$ (задан ниже в этом тексте) для того, чтобы создавать tr_k , чтобы достигать более точных значений E_t и E_r .

Вычислим:

$$f_k = \frac{E_{ok}}{E_{tk}}$$

если $E_{tk} > 0$, иначе $f_k = 0$.

С помощью:

$$E_k = \sqrt{f_k E_{rk}}$$

далее вычисляется более стабильная версия E_k , поскольку вычисление E_k с помощью MDCT-значений искажается только посредством того факта, что MDCT-значения не подчиняются теореме Парсеваля, и следовательно, они не отражают информацию полной энергии спектральных значений. \tilde{E}_k вычисляется так, как описано выше.

Как отмечено выше, для коротких блоков предположим, что кодер решает группировать num_window_group наборов коэффициентов масштабирования. Как описано выше, w_l обозначает l -ый поднабор w , где l обозначает индекс группы окон кодирования со взвешиванием, $0 \leq l < num_window_group$.

С другой стороны, может вычисляться альтернативная версия, указанная выше для того, чтобы вычислять более стабильную версию $E_{k,l}$. Если задать $\hat{c} := (\hat{x}_r, \hat{x}_i) \in \mathbb{C}^N$, причем $\hat{x}_r \in \mathbb{R}^N$ является MDCT-преобразованным, а $\hat{x}_i \in \mathbb{R}^N$ является MDST-преобразованным кодированным со взвешиванием аудиосигналом с длиной $2N$, вычисляем:

$$E_{okl} = \frac{1}{|w_l|} \sum_{l \in w_l} \frac{1}{|scb_k|} \sum_{i \in scb_k} \hat{c}_{i,l}^2$$

Аналогично вычисляем:

$$E_{tkl} = \frac{1}{|w_l|} \sum_{l \in w_l} \frac{1}{|scb_k|} \sum_{i \in tr_k} \hat{c}_{i,l}^2, \quad E_{rkl} = \frac{1}{|w_l|} \sum_{l \in w_l} \frac{1}{|scb_k|} \sum_{i \in tr_k} \hat{x}_{r,l}^2$$

и продолжаем для коэффициента $f_{k,l}$:

$$f_{k,l} = \frac{E_{okl}}{E_{tkl}},$$

который используется для того, чтобы регулировать ранее вычисленный $E_{rk,l}$:

$$E_{k,l} = \sqrt{f_{k,l} E_{rk,l}}$$

$\tilde{E}_{k,l}$ вычисляется так, как описано выше.

Процедура не только использования энергии полосы частот восстановления, извлекаемой либо из комплексной полосы частот восстановления, либо из MDCT-значений, но также и использования информации энергии из исходного диапазона обеспечивает усовершенствованное восстановление энергии.

В частности, модуль 1006 вычисления параметров выполнен с возможностью вычислять информацию энергии для полосы частот восстановления с использованием информации относительно энергии полосы частот восстановления и дополнительно с использованием информации относительно энергии исходного диапазона, которая должна быть использована для восстановления полосы частот восстановления.

Кроме того, модуль 1006 вычисления параметров выполнен с возможностью вычислять информацию (E_{ok}) энергии для полосы частот восстановления комплексного спектра исходного сигнала, вычислять дополнительную информацию (E_{rk}) энергии для исходного диапазона действительнзначной части комплексного спектра исходного сигнала, который должен использоваться для восстановления полосы частот восстановления, и при этом модуль вычисления параметров выполнен с возможностью вычислять информацию энергии для полосы частот восстановления с использованием информации (E_{ok}) энергии и дополнительной информации (E_{rk}) энергии.

Кроме того, модуль 1006 вычисления параметров выполнен с возможностью определения первой информации (E_{ok}) энергии относительно подлежащей восстановлению полосы частот коэффициентов масштабирования комплексного спектра исходного сигнала, с возможностью определения второй информации (E_{tk}) энергии относительно исходного диапазона комплексного спектра исходного сигнала, который должен использоваться для восстановления подлежащей восстановлению полосы частот коэффициентов масштабирования, с возможностью определения третьей информации (E_{rk}) энергии относительно исходного диапазона действительнзначной части комплексного спектра исходного сигнала, который должен использоваться для восстановления подлежащей восстановлению полосы частот коэффициентов масштабирования, с возможностью определения информации взвешивания на основе взаимосвязи, по меньшей мере, между двумя из первой информации энергии, второй информации энергии и третьей информации энергии, и с возможностью взвешивания одной из первой информации энергии и третьей информации энергии с использованием информации взвешивания, чтобы получать информацию взвешенной энергии, и с возможностью использования информации взвешенной энергии в качестве информации энергии для полосы частот восстановления.

Примеры для вычислений являются следующими, но множество других могут быть очевидными для специалистов в данной области техники с учетом вышеизложенного общего принципа:

A)

$$f_k = E_{ok} / E_{tk};$$

$$E_k = \sqrt{f_k * E_{rk}};$$

B)

$$f_k = E_{tk} / E_{ok};$$

$$E_k = \sqrt{(1/f_k) * E_{rk}};$$

C)

$$f_k = E_{rk} / E_{tk};$$

$$E_k = \sqrt{f_k * E_{ok}};$$

D)

$$f_k = E_{tk} / E_{rk};$$

$$E_k = \sqrt{(1/f_k) * E_{ok}}$$

Все эти примеры подтверждают тот факт, что хотя только действительные MDCT-значения обрабатываются на стороне декодера, фактическое вычисление (вследствие суммирования с перекрытием) проводится из процедуры подавления наложения спектров во временной области, неявно выполняемой с использованием комплексных чисел. Тем не менее, в частности, определение 918 информации энергии мозаичных фрагментов дополнительных спектральных частей 922, 923 полосы 920 частот восстановления для значений частоты, отличающихся от первой спектральной части 921, имеющей частоты в полосе 920 частот восстановления, основывается на действительных MDCT-значениях. Следовательно, информация энергии, передаваемая в декодер, типично меньше информации E_{ok} энергии относительно полосы частот восстановления комплексного спектра исходного сигнала. Например, для вышеприведенного случая С, это означает то, что коэффициент f_k (информация взвешивания) меньше 1.

На стороне декодера, если инструментальное IGF-средство передается в служебных сигналах как активированное, передаваемые значения \hat{E}_k получаются из потока битов и должны быть деквантованы следующим образом:

$$E_k = 2^{\frac{1}{4}\hat{E}_k},$$

для всех $k = igfStartSfb, 1 + igfStartSfb, 2 + igfStartSfb, \dots, igfEndSfb$.

Декодер деквантует передаваемые MDCT-значения до $x \in \mathbb{R}^N$ и вычисляет оставшуюся энергию выживания:

$$sE_k := \sum_{i \in scb_k} x_i^2,$$

где k находится в диапазоне, заданном выше.

Обозначим $\overline{scb_k} = \{i | i \in scb_k \wedge x_i = 0\}$. Этот набор содержит все индексы полосы scb_k частот коэффициентов масштабирования, которые квантованы до нуля посредством кодера.

IGF-способ получения подполос частот (не описан в данном документе) используется для того, чтобы заполнять интервалы отсутствия сигнала в спектре, получающиеся в результате приблизительного квантования спектральных MDCT-значений на стороне кодера посредством использования ненулевых значений передаваемого MDCT. x должно дополнительно содержать значения, которые заменяют все предыдущие обнуленные значения. Энергия мозаичных фрагментов вычисляется следующим образом:

$$tE_k := \sum_{i \in \overline{scb_k}} x_i^2,$$

где k находится в диапазоне, заданном выше.

Энергия, недостающая в полосе частот восстановления, вычисляется следующим образом:

$$mE_k := |scb_k| E_k^2 - sE_k$$

Кроме того, коэффициент усиления для регулирования получается следующим образом:

$$g := \begin{cases} \sqrt{\frac{mE_k}{tE_k}} & \text{if } (mE_k > 0 \wedge tE_k > 0), \\ 0 & \text{else} \end{cases}$$

где:

$$g' = \min(g, 10)$$

Регулирование спектральной огибающей с использованием коэффициента усиления следующее:

$$x_i := g' x_i,$$

для всех $i \in \overline{scb_k}$, и k находится в диапазоне, заданном выше.

Это повторно придает спектральной огибающей x определенную форму исходной спектральной огибающей \hat{x} .

В последовательности коротких окон кодирования со взвешиванием, все вычисления, приведенные выше, в принципе остаются неизменными, но учитывается группировка полос частот коэффициентов масштабирования. Обозначим в качестве $E_{k,l}$ деквантованные сгруппированные значения энергии, полученные из потока битов.

Вычислим:

$$sE_{k,l} := \frac{1}{|w_l|} \sum_{j \in w_l} \sum_{i \in scb_{j,k}} x_{j,i}^2$$

и:

$$pE_{k,l} := \frac{1}{|w_l|} \sum_{j \in w_l} \sum_{i \in scb_{j,k}} x_{j,i}^2$$

Индекс j описывает индекс окна кодирования со взвешиванием последовательности коротких блоков.

Вычислим:

$$mE_{k,l} := |scb_k| E_{k,l}^2 - sE_{k,l}$$

и:

$$g := \begin{cases} \sqrt{\frac{mE_{k,l}}{pE_{k,l}}} & \text{if } (mE_{k,l} > 0 \wedge pE_{k,l} > 0), \\ 0 & \text{else} \end{cases}$$

где:

$$g' = \min(g, 10)$$

Применим:

$$x_{j,i} := g' x_{j,i},$$

для всех $i \in \overline{scb_{k,l}}$.

Для вариантов применения с низкой скоростью передачи битов возможна попарная группировка значений E_k без серьезной потери точности. Этот способ применяется только для длинных блоков:

$$E_{k \gg 1} = \sqrt{\frac{1}{|\overline{scb_k} \cup \overline{scb_{k+1}}|} \sum_{i \in \overline{scb_k} \cup \overline{scb_{k+1}}} \hat{x}_i^2},$$

где $k = igfStartSfb, 2 + igfStartSfb, 4 + igfStartSfb, \dots, igfEndSfb$.

С другой стороны, после квантования, все значения $E_{k \gg 1}$ передаются в декодер.

Фиг. 9а иллюстрирует устройство для декодирования кодированного аудиосигнала, содержащего кодированное представление первого набора первых спектральных частей и кодированное представление параметрических данных, указывающих спектральные

энергии для второго набора вторых спектральных частей. Первый набор первых спектральных частей указывается на 901a на фиг. 9a, а кодированное представление параметрических данных указывается на 901b на фиг. 9a. Аудиодекодер 900 предоставляется для декодирования кодированного представления 901a первого набора
 5 первых спектральных частей для того, чтобы получать декодированный первый набор первых спектральных частей 904, и для декодирования кодированного представления параметрических данных для того, чтобы получать декодированные параметрические данные 902 для второго набора вторых спектральных частей, указывающих отдельные энергии для отдельных полос частот восстановления, причем вторые спектральные
 10 части расположены в полосах частот восстановления. Кроме того, модуль 906 повторного формирования частоты предоставляется для восстановления спектральных значений полосы частот восстановления, содержащей вторую спектральную часть. Модуль 906 повторного формирования частоты использует первую спектральную часть первого набора первых спектральных частей и информацию отдельных энергий
 15 для полосы частот восстановления, при этом полоса частот восстановления содержит первую спектральную часть и вторую спектральную часть. Модуль 906 повторного формирования частоты содержит модуль 912 вычисления для определения информации энергии выживания, содержащей накопленную энергию первой спектральной части, имеющей частоты в полосе частот восстановления. Кроме того, модуль 906 повторного
 20 формирования частоты содержит модуль 918 вычисления для определения информации энергии мозаичных фрагментов дополнительных спектральных частей полосы частот восстановления и для значений частоты, отличающихся от первой спектральной части, причем эти значения частоты имеют частоты в полосе частот восстановления, при этом дополнительные спектральные части должны формироваться посредством повторного
 25 формирования частоты с использованием первой спектральной части, отличающейся от первой спектральной части в полосе частот восстановления.

Модуль 906 повторного формирования частоты дополнительно содержит модуль 914 вычисления для недостающей энергии в полосе частот восстановления, и модуль 914 вычисления работает с использованием отдельной энергии для полосы частот
 30 восстановления и энергии выживания, сформированной посредством блока 912. Кроме того, модуль 906 повторного формирования частоты содержит модуль 916 регулирования спектральной огибающей для регулирования дополнительных спектральных частей в полосе частот восстановления на основе информации недостающей энергии и информации энергии мозаичных фрагментов, сформированной
 35 посредством блока 918.

Следует обратиться к фиг. 9с, иллюстрирующий определенную полосу 920 частот восстановления. Полоса частот восстановления содержит первую спектральную часть в полосе частот восстановления, к примеру, первую спектральную часть 306 на фиг. 3а, схематично проиллюстрированную на 921. Кроме того, остальные спектральные
 40 значения в полосе 920 частот восстановления должны формироваться с использованием исходной области, например, из полосы 1, 2, 3 частот коэффициентов масштабирования ниже начальной частоты 309 интеллектуального заполнения интервалов по фиг. 3а. Модуль 906 повторного формирования частоты выполнен с возможностью формирования необработанных спектральных значений для вторых спектральных
 45 частей 922 и 923. Затем коэффициент g усиления вычисляется так, как проиллюстрировано на фиг. 9с, чтобы, в завершение, регулировать необработанные спектральные значения в полосах частот 922, 923 так, чтобы получать восстановленные и отрегулированные вторые спектральные части в полосе 920 частот восстановления,

которые в данный момент имеют спектральное разрешение, т.е. линейное расстояние, идентичное спектральному разрешению (линейному расстоянию) первой спектральной части 921. Важно понимать, что первая спектральная часть в полосе частот восстановления, проиллюстрированная на 921 на фиг. 9с, декодируется посредством аудиодекодера 900 и не затрагивается посредством регулирования огибающей, выполняемого посредством блока 916 по фиг. 9b. В отличие от этого, первая спектральная часть в полосе частот восстановления, указываемая на 921, остается как есть, поскольку эта первая спектральная часть выводится посредством полноскоростного (с полной полосой пропускания) аудиодекодера 900 через линию 904.

Далее поясняется конкретный пример с действительными числами. Оставшаяся энергия выживания, вычисленная посредством блока 912, составляет, например, пять энергетических единиц, и эта энергия представляет собой энергию примерно указываемых четырех спектральных линий в первой спектральной части 921.

Кроме того, значение E3 энергии для полосы частот восстановления, соответствующей полосе 6 частот коэффициентов масштабирования по фиг. 3b или фиг. 3a, равно 10 единицам. Важно, что значение энергии содержит не только энергию спектральных частей 922, 923, но и полную энергию полосы 920 частот восстановления, вычисленную на стороне кодера, т.е. до выполнения спектрального анализа с использованием, например, маски тональности. Следовательно, десять энергетических единиц покрывают первую и вторую спектральные части в полосе частот восстановления. Далее предполагается, что энергия данных исходного диапазона для блоков 922, 923 или для необработанных данных целевого диапазона для блока 922, 923 равна восьми энергетическим единицам. Таким образом, вычисляется недостающая энергия в пять единиц.

На основе недостающей энергии, деленной на энергию tEk мозаичных фрагментов, вычисляется коэффициент усиления в 0,79. После этого необработанные спектральные линии для вторых спектральных частей 922, 923 умножаются на вычисленный коэффициент усиления. Таким образом, регулируются только спектральные значения для вторых спектральных частей 922, 923, и спектральные линии для первой спектральной части 921 не затрагиваются посредством этого регулирования огибающей. После умножения необработанных спектральных значений для вторых спектральных частей 922, 923, вычислена полная полоса частот восстановления, состоящая из первых спектральных частей в полосе частот восстановления и состоящая из спектральных линий во вторых спектральных частях 922, 923 в полосе 920 частот восстановления.

Предпочтительно, исходный диапазон для формирования необработанных спектральных данных в полосах частот 922, 923, относительно частоты, ниже начальной IGF-частоты 309, а полоса 920 частот восстановления выше начальной IGF-частоты 309.

Кроме того, предпочтительно, чтобы границы полос частот восстановления совпадали с границами полос частот коэффициентов масштабирования. Таким образом, полоса частот восстановления имеет, в одном варианте осуществления, размер соответствующих полос частот коэффициентов масштабирования базового аудиодекодера, либо имеет такой размер, что когда применяется энергетическое спаривание, значение энергии для полосы частот восстановления предоставляет энергию в две или более высокое целое число полос частот коэффициентов масштабирования. Таким образом, когда предполагается, что аккумулярование энергии выполняется для полосы 4 частот коэффициентов масштабирования, полосы 5 частот коэффициентов

масштабирования и полосы 6 частот коэффициентов масштабирования, в таком случае нижняя частотная граница полосы 920 частот восстановления равна нижней границе полосы 4 частот коэффициентов масштабирования, а верхняя частотная граница полосы 920 частот восстановления совпадает с верхней границей полосы 6 частот коэффициентов масштабирования.

Далее поясняется фиг. 9d для того, чтобы показывать дополнительные функциональности декодера по фиг. 9a. Аудиодекодер 900 принимает деквантованные спектральные значения, соответствующие первым спектральным частям первого набора спектральных частей, и дополнительно, коэффициенты масштабирования для полос частот коэффициентов масштабирования, к примеру, проиллюстрированных на фиг. 3b предоставляются в блок 940 обратного масштабирования. Блок 940 обратного масштабирования предоставляет все первые наборы первых спектральных частей ниже начальной IGF-частоты 309 по фиг. 3a и, дополнительно, первые спектральные части выше начальной IGF-частоты, т.е. первые спектральные части 304, 305, 306, 307 по фиг. 3a, которые находятся в полосе частот восстановления, как проиллюстрировано на 941 на фиг. 9d. Кроме того, первые спектральные части в исходной полосе частот, используемые для заполнения частотными мозаичными фрагментами в полосе частот восстановления, предоставляются в модуль 942 регулирования/вычисления огибающей, и этот блок дополнительно принимает информацию энергии для полосы частот восстановления, предоставленную в качестве параметрической вспомогательной информации в кодированном аудиосигнале, как проиллюстрировано на 943 на фиг. 9d. Затем модуль 942 регулирования/вычисления огибающей предоставляет функциональности по фиг. 9b и 9c и, в завершение, выводит отрегулированные спектральные значения для вторых спектральных частей в полосе частот восстановления. Эти отрегулированные спектральные значения 922, 923 для вторых спектральных частей в полосе частот восстановления и первых спектральных частей 921 в полосе частот восстановления указывают то, что линия 941 на фиг. 9d объединенно представляет полное спектральное представление полосы частот восстановления.

Далее следует обратиться к фиг. 10a в 10b для пояснения предпочтительных вариантов осуществления аудиокодера для кодирования аудиосигнала, чтобы предоставлять или формировать кодированный аудиосигнал. Кодер содержит временно-спектральный преобразователь 1002, подающий спектральный анализатор 1004, и спектральный анализатор 1004 соединяется с модулем 1006 вычисления параметров, с одной стороны, и с аудиокодером 1008, с другой стороны. Аудиокодер 1008 предоставляет кодированное представление первого набора первых спектральных частей и не покрывает второй набор вторых спектральных частей. С другой стороны, модуль 1006 вычисления параметров предоставляет информацию энергии для полосы частот восстановления, покрывающей первые и вторые спектральные части. Кроме того, аудиокодер 1008 выполнен с возможностью формирования первого кодированного представления первого набора первых спектральных частей, имеющих первое спектральное разрешение, при этом аудиокодер 1008 предоставляет коэффициенты масштабирования для всех полос частот спектрального представления, сформированного посредством блока 1002. Кроме того, как проиллюстрировано на фиг. 3b, кодер предоставляет информацию энергии, по меньшей мере, для полос частот восстановления, расположенных, относительно частоты, выше начальной IGF-частоты 309, как проиллюстрировано на фиг. 3a. Таким образом, для полос частот восстановления, предпочтительно совпадающих с полосами частот коэффициентов масштабирования или с группами полос частот коэффициентов масштабирования, предоставляются два значения, т.е.

соответствующий коэффициент масштабирования из аудиокодера 1008 и, дополнительно, информация энергии, выводимая посредством модуля 1006 вычисления параметров.

Аудиокодер предпочтительно имеет полосы частот коэффициентов масштабирования с различными полосами пропускания частот, т.е. с различным количеством спектральных значений. Следовательно, параметрический модуль вычисления содержит нормализатор 1012 для нормализации энергий для различной полосы пропускания относительно полосы пропускания конкретной полосы частот восстановления. С этой целью, нормализатор 1012 принимает, в качестве вводов, энергию в полосе частот и количество спектральных значений в полосе частот, и нормализатор 1012 затем выводит нормализованную энергию в расчете на полосу частот восстановления/коэффициентов масштабирования.

Кроме того, параметрический модуль 1006а вычисления по фиг. 10а содержит модуль вычисления значений энергии, принимающей управляющую информацию из базового кодера или аудиокодера 1008, как проиллюстрировано посредством линии 1007 на фиг. 10а. Эта управляющая информация может содержать информацию относительно длинных/коротких блоков, используемых посредством аудиокодера, и/или информацию группировки. Следовательно, в то время как информация относительно длинных/коротких блоков и информация группировки относительно коротких окон кодирования со взвешиванием связаны с "временной" группировкой, информация группировки дополнительно может означать спектральную группировку, т.е. группировку двух полос частот коэффициентов масштабирования в одну полосу частот восстановления. Следовательно, модуль 1014 вычисления значений энергии выводит одно значение энергии для каждой сгруппированной полосы частот, покрывающей первую и вторую спектральную часть, когда только спектральные части сгруппированы.

Фиг. 10d иллюстрирует дополнительный вариант осуществления для реализации спектральной группировки. С этой целью, блок 1016 выполнен с возможностью вычисления значений энергии для двух смежных полос частот. Затем в блоке 1018 сравниваются значения энергии для смежных полос частот, и когда значения энергии сильно не отличаются или отличаются в меньшей степени, чем задано, например, посредством порогового значения, то формируется одно (нормализованное) значение для обеих полос частот, как указано в блоке 1020. Как проиллюстрировано посредством линии 1019, блок 1018 может обходиться. Кроме того, формирование одного значения для двух или более полос частот, выполняемое посредством блока 1020, может управляться посредством управления 1024 скоростью передачи битов кодера. Таким образом, когда скорость передачи битов должна уменьшаться, управление 1024 кодированной скоростью передачи битов управляет блоком 1020 таким образом, чтобы формировать одно нормализованное значение для двух или более полос частот, даже если сравнение в блоке 1018 не разрешено для того, чтобы группировать значения информации энергии.

В случае если аудиокодер выполняет группировку двух или более коротких окон кодирования со взвешиванием, эта группировка также применяется для информации энергии. Когда базовый кодер выполняет группировку двух или более коротких блоков, то для двух или более блоков вычисляется и передается только один набор коэффициентов масштабирования. На стороне декодера, аудиодекодер затем применяет идентичный набор коэффициентов масштабирования для обоих сгруппированных окон кодирования со взвешиванием.

Относительно вычисления информации энергии, спектральные значения в полосе частот восстановления накапливаются за два или более коротких окна кодирования

со взвешиванием. Другими словами, это означает то, что спектральные значения в определенной полосе частот восстановления для короткого блока и для последующего короткого блока накапливаются вместе, и только одно значение информации энергии передается для этой полосы частот восстановления, покрывающей два коротких блока.

5 Затем на стороне декодера, регулирование огибающей, поясненное относительно фиг. 9a на 9d, выполняется не по отдельности для каждого короткого блока, а совместно для набора сгруппированных коротких окон кодирования со взвешиванием.

После этого снова применяется соответствующая нормализация, так что даже если выполнена группировка по частоте или группировка во времени, нормализация легко

10 обеспечивает то, что для вычисления информации значений энергии на стороне декодера, должно быть известно только значение информации энергии, с одной стороны, и количество спектральных линий в полосе частот восстановления или в наборе сгруппированных полос частот восстановления.

В BWE-схемах предшествующего уровня техники, восстановление спектральной HF-

15 области выше данной так называемой частоты разделения зачастую основано на спектральном наложении. Типично, HF-область состоит из нескольких смежных наложений, и каждое из этих наложений получается из полосовых (BP) областей LF-спектра ниже данной частоты разделения. В представлении на основе гребенки фильтров сигнала такие системы копируют набор смежных подполосных коэффициентов из LF-

20 спектра в целевую область. Границы выбранных наборов типично являются системно-зависимыми и не являются зависимыми от сигнала. Для некоторого содержимого сигнала, этот статический выбор наложений может приводить к неприятному тембру и окраске восстановленного сигнала.

Другие подходы передают LF-сигнал в HF через сигнально-адаптивную модуляцию

25 с одной боковой полосой (SSB). Такие подходы имеют высокую вычислительную сложность по сравнению с [1], поскольку они работают на высокой частоте дискретизации в выборках временной области. Кроме того, наложение может становиться нестабильным, особенно для нетональных сигналов (например, невокализованной речи), и в силу этого сигнально-адаптивное наложение

30 предшествующего уровня техники может вводить ухудшения в сигнал.

Изобретаемый подход называется "интеллектуальным заполнением интервалов (IGF)", и в предпочтительной конфигурации он применяется в BWE-системе на основе частотно-временного преобразования, такого как, например, модифицированное дискретное косинусное преобразование (MDCT). Тем не менее, идеи изобретения

35 являются, в общем, применимыми, например, аналогично в системе на основе гребенки квадратурных зеркальных фильтров (QMF).

Преимущество IGF-конфигурации на основе MDCT заключается в прозрачной интеграции в аудиокодеры на основе MDCT, например, для усовершенствованного кодирования аудио (AAC) на основе MPEG. Совместное использование идентичного

40 преобразования для кодирования аудио формы сигнала и для BWE значительно снижает общую вычислительную сложность для аудиокодека.

Кроме того, изобретение предоставляет решение для внутренне присущих проблем стабильности, выявленных в схемах адаптивного наложения предшествующего уровня техники.

Предложенная система основана на таком наблюдении, что для некоторых сигналов, ненаправленный выбор наложений может приводить к изменениям тембра и окраскам сигналов. Если сигнал, который является тональным в исходной спектральной области (SSR), но является шумоподобным в целевой спектральной области (STR), наложение

шумоподобной STR посредством тональной SSR может приводить к неестественному тембру. Тембр сигнала также может изменяться, поскольку тональная структура сигнала может расстраиваться или даже уничтожаться посредством процесса наложения.

Предложенная IGF-система выполняет интеллектуальный выбор мозаичных фрагментов с использованием взаимной корреляции в качестве меры подобия между конкретной SSR и конкретной STR. Взаимная корреляция двух сигналов предоставляет меру подобия этих сигналов, а также запаздывание максимальной корреляции и ее знак. Следовательно, подход с выбором мозаичных фрагментов на основе корреляции также может использоваться для того, чтобы точно регулировать спектральное смещение скопированного спектра так, что он становится максимально близким к исходной спектральной структуре.

Фундаментальный вклад предложенной системы заключается в выборе подходящей меры подобия, а также технологий для того, чтобы стабилизировать процесс выбора мозаичных фрагментов. Предложенная технология предоставляет оптимальный баланс между мгновенной сигнальной адаптацией и в то же время временной стабильностью. Обеспечение временной стабильности является особенно важным для сигналов, которые имеют незначительное подобие SSR и STR, и следовательно, демонстрируют низкие значения взаимной корреляции, либо если используются меры подобия, которые являются неоднозначными. В таких случаях, стабилизация предотвращает псевдослучайное поведение адаптивного выбора мозаичных фрагментов.

Например, класс сигналов, который зачастую приводит к проблемам для BWE предшествующего уровня техники, отличается посредством различной концентрации энергии в произвольных спектральных областях, как показано на фиг. 12а (слева). Хотя доступны способы для того, чтобы регулировать спектральную огибающую и тональность восстановленного спектра в целевой области, для некоторых сигналов эти способы не позволяют хорошо сохранять тембр, как показано на фиг. 12а (справа). В примере, показанном на фиг. 12а, абсолютная величина спектра в целевой области исходного сигнала выше так называемой частоты $f_{\text{crossover}}$ разделения (фиг. 12а, слева) снижается практически линейно. Напротив, в восстановленном спектре (фиг. 12а, справа), присутствует отдельный набор падений и пиков, который воспринимается как артефакт окрашивания тембра.

Важный этап нового подхода заключается в том, чтобы задавать набор мозаичных фрагментов, из числа которых может осуществляться последующий выбор на основе подобия. Во-первых, границы мозаичных фрагментов как исходной области, так и целевой области должны задаваться друг в соответствии с другом. Следовательно, целевая область между начальной IGF-частотой f_{IGFstart} базового кодера и наибольшей доступной частотой f_{IGFstop} разделена на произвольное целое число n_{Tar} мозаичных фрагментов, каждый из которых имеет индивидуальный предварительно заданный размер. Затем для каждого целевого мозаичного фрагмента $\text{tar}[\text{idx_tar}]$, формируется набор $\text{src}[\text{idx_src}]$ исходных мозаичных фрагментов одинакового размера. Посредством этого, определяется базовая степень свободы IGF-системы. Общее число n_{Src} исходных мозаичных фрагментов определяется посредством полосы пропускания исходной области:

$$bw_{\text{src}} = (f_{\text{IGFstart}} - f_{\text{IGFmin}}),$$

где f_{IGFmin} является наименьшей доступной частотой для выбора мозаичных фрагментов, так что целое число n_{Src} исходных мозаичных фрагментов подходит для

bw_{src} . Минимальное число исходных мозаичных фрагментов равно 0.

Чтобы дополнительно повышать степень свободы для выбора и регулирования, исходные мозаичные фрагменты могут задаваться таким образом, что они перекрывают друг друга на коэффициент перекрытия между 0 и 1, где 0 означает отсутствие
 5 перекрытия, а 1 означает 100%-ое перекрытие. Случай 100%-ого перекрытия подразумевает то, что доступен только один или ни одного исходного мозаичного фрагмента.

Фиг. 12b показывает пример границ мозаичных фрагментов из набора мозаичных фрагментов. В этом случае, все целевые мозаичные фрагменты коррелируются с каждым
 10 из исходных мозаичных фрагментов. В этом примере, исходные мозаичные фрагменты перекрываются на 50%.

Для целевого мозаичного фрагмента, взаимная корреляция вычисляется с различными исходными мозаичными фрагментами при запаздываниях вплоть до элементов $xcorr_maxLag$ выборки. Для данного целевого мозаичного фрагмента idx_tar и исходного
 15 мозаичного фрагмента idx_src , $xcorr_val[idx_tar][idx_src]$ обеспечивает максимальное значение абсолютной взаимной корреляции между мозаичными фрагментами, тогда как $xcorr_lag[idx_tar][idx_src]$ обеспечивает запаздывание, с которым возникает этот максимум, а $xcorr_sign[idx_tar][idx_src]$ обеспечивает знак перекрестной корреляции при
 20 $xcorr_lag[idx_tar][idx_src]$.

Параметр $xcorr_lag$ используется для того, чтобы управлять близостью совпадения между исходными и целевыми мозаичными фрагментами. Этот параметр приводит к уменьшенным артефактам и помогает лучше сохранять тембр и цвет сигнала.

В некоторых сценариях может возникать такая ситуация, что размер конкретного
 25 целевого мозаичного фрагмента больше размера доступных исходных мозаичных фрагментов. В этом случае, доступный исходный мозаичный фрагмент повторяется так часто, как требуется для того, чтобы полностью заполнять конкретный целевой мозаичный фрагмент. По-прежнему можно выполнять взаимную корреляцию между крупным целевым мозаичным фрагментом и меньшим исходным мозаичным фрагментом
 30 для того, чтобы получать оптимальную позицию исходного мозаичного фрагмента в целевом мозаичном фрагменте с точки зрения запаздывания $xcorr_lag$ и знака $xcorr_sign$ взаимной корреляции.

Взаимная корреляция необработанных спектральных мозаичных фрагментов и исходного сигнала может не представлять собой наиболее подходящую меру подобия,
 35 применяемую к спектрам звука с сильной формантной структурой. Отбеливание спектра удаляет приблизительную информацию огибающей и за счет этого подчеркивает точную спектральную структуру, которая представляет главный интерес для оценки подобия мозаичных фрагментов. Отбеливание также помогает при простом формировании огибающей STR в декодере для областей, обработанных посредством IGF.
 40 Следовательно, необязательно, мозаичный фрагмент и исходный сигнал отбеливаются до вычисления взаимной корреляции.

В других конфигурациях, только мозаичный фрагмент отбеливается с использованием предварительно заданной процедуры. Передаваемый флаг "отбеливания" указывает
 45 декодеру, что идентичный предварительно заданный процесс отбеливания должен применяться к мозаичному фрагменту в IGF.

Для отбеливания сигнала сначала вычисляется оценка спектральной огибающей. Затем MDCT-спектр делится на спектральную огибающую. Оценка спектральной огибающей может оцениваться для MDCT-спектра, энергий MDCT-спектра,

комплексного спектра мощности на основе MDCT или оценок спектра мощности. Сигнал, для которого оценивается огибающая, далее называется "базовым сигналом".

Огибающие, вычисленные для комплексного спектра мощности на основе MDCT или оценок спектра мощности в качестве базового сигнала, имеют преимущество отсутствия временной флуктуации для тональных компонентов.

Если базовый сигнал находится в энергетической области, MDCT-спектр должен быть поделен на квадратный корень огибающей, чтобы корректно отбеливать сигнал.

Предусмотрены различные способы вычисления огибающей:

- преобразование базового сигнала с помощью дискретного косинусного преобразования (DCT), сохраняющего только более низкие DCT-коэффициенты (задающего самый верхний равным нулю), и затем вычисление обратного DCT,
- вычисление спектральной огибающей набора коэффициентов линейного прогнозирования (LPC), вычисленных для аудиокадра временной области,
- фильтрация базового сигнала с помощью фильтра нижних частот.

Предпочтительно, выбирается последний подход. Для вариантов применения, которые требуют низкой вычислительной сложности, некоторое упрощение может вноситься для отбеливания MDCT-спектра: Сначала огибающая вычисляется посредством скользящего среднего. Это требует только двух циклов процессора в расчете на элемент MDCT-выборки. Затем во избежание вычисления деления и квадратного корня, спектральная огибающая аппроксимирована посредством 2^n , где n является целочисленным логарифмом огибающей. В этой области, операция вычисления квадратного корня становится просто операцией сдвига, и кроме того, деление на огибающую может выполняться посредством другой операции сдвига.

После вычисления корреляции каждого исходного мозаичного фрагмента с каждым целевым мозаичным фрагментом для всех $nTar$ целевых мозаичных фрагментов, исходный мозаичный фрагмент с наибольшей корреляцией выбирается для его замены. Для обеспечения наилучшего совпадения с исходной спектральной структурой, запаздывание корреляции используется для того, чтобы модулировать реплицируемый спектр посредством целого числа элементов выборки преобразования. В случае нечетных запаздываний мозаичный фрагмент дополнительно модулирован через умножение на переменную временную последовательность из $-1/1$ для того, чтобы компенсировать представление с обратной частотой каждой второй полосы частот в MDCT.

Фиг. 12с показывает пример корреляции между исходным мозаичным фрагментом и целевым мозаичным фрагментом. В этом примере, запаздывание корреляции равно 5, так что исходный мозаичный фрагмент должен быть модулирован посредством 5 элементов выборки в направлении элементов выборки верхних частот на стадии перезаписи из BWE-алгоритма. Помимо этого, знак мозаичного фрагмента должен меняться на противоположный, поскольку значение максимальной корреляции является отрицательным, и дополнительная модуляция, как описано выше, учитывает нечетное запаздывание.

Таким образом, общий объем вспомогательной информации, которую следует передавать из кодера в декодер, может состоять из следующих данных:

$tileNum[nTar]$: индекс выбранного исходного мозаичного фрагмента в расчете на целевой мозаичный фрагмент

$tileSign[nTar]$: знак целевого мозаичного фрагмента

$tileMod[nTar]$: запаздывание корреляции в расчете на целевой мозаичный фрагмент

Отсечение и стабилизация мозаичных фрагментов является важным этапом в IGF.

Его необходимость и преимущества поясняются с помощью примера при допущении стационарного тонального аудиосигнала, такого как, например, стабильное звучание камертона. Логика подсказывает, что наименьшее количество артефактов вводится, если для данной целевой области исходные мозаичные фрагменты всегда выбираются из идентичной исходной области между кадрами. Даже если сигнал предположительно является стационарным, это условие не выполняется полностью в каждом кадре, поскольку мера подобия (например, корреляция) другой в равной степени подобной исходной области может доминировать над результатом оценки подобия (например, взаимной корреляции). Это приводит к тому, что $\text{tileNum}[n\text{Tar}]$ между смежными кадрами колеблется между двумя или тремя практически идентичными вариантами выбора. Это может представлять собой источник раздражающего музыкального шумоподобного артефакта.

Чтобы исключать этот тип артефактов, набор исходных мозаичных фрагментов должен быть отсечен таким образом, что оставшиеся элементы исходного набора являются максимально несходными. Это достигается для набора исходных мозаичных фрагментов:

$$S=\{s_1, s_2, \dots, s_n\}$$

следующим образом. Для любого исходного мозаичного фрагмента s_i , он коррелируется со всеми другими исходными мозаичными фрагментами, находя наилучшую корреляцию между s_i и s_j и сохраняя ее в матрице S_x . Здесь $S_x[i][j]$ содержит максимальное значение абсолютной взаимной корреляции между s_i и s_j . Суммирование матрицы S_x по столбцам дает сумму перекрестных корреляций исходного мозаичного фрагмента s_i со всеми другими исходными мозаичными фрагментами T .

$$T[i]=S_x[i][1]+S_x[i][2]+\dots+S_x[i][n]$$

Здесь T представляет меру того, насколько сильно источник является подобным другим исходным мозаичным фрагментам. Если, для любого исходного мозаичного фрагмента i :

$T >$ пороговое значение,

то исходный мозаичный фрагмент i может быть отброшен из набора потенциальных источников, поскольку он имеет высокую корреляцию с другими источниками.

Мозаичный фрагмент с наименьшей корреляцией из набора мозаичных фрагментов, которые удовлетворяют условию в уравнении 1, выбран в качестве характерного мозаичного фрагмента для этого поднабора. Таким образом, обеспечивается то, что исходные мозаичные фрагменты являются максимально несходными друг с другом.

Способ отсечения мозаичных фрагментов также включает в себе запоминающее устройство для набора отсеченных мозаичных фрагментов, используемого в предшествующем кадре. Мозаичные фрагменты, которые являются активными в предыдущем кадре, также сохраняются в следующем кадре, если существуют альтернативные возможные варианты для отсечения.

Пусть исходные мозаичные фрагменты s_3, s_4 и s_5 являются активными из мозаичных фрагментов $\{s_1, s_2, \dots, s_5\}$ в кадре k , затем в кадре $k+1$, даже если исходные мозаичные фрагменты s_1, s_3 и s_2 конкурируют за отсечение, причем s_3 является максимально коррелированным с другими, s_3 сохраняется, поскольку он представляет собой полезный исходный мозаичный фрагмент в предыдущем кадре, и в силу этого его сохранение в наборе исходных мозаичных фрагментов является полезным для принудительной активации временной непрерывности при выборе мозаичных фрагментов. Этот способ предпочтительно применяется, если взаимная корреляция между исходным i и целевым j , представленная как $T_x[i][j]$, является высокой.

Дополнительный способ для стабилизации мозаичных фрагментов состоит в том, чтобы сохранять порядок мозаичных фрагментов из предыдущего кадра $k-1$, если ни один из исходных мозаичных фрагментов в текущем кадре k не коррелируется хорошо с целевыми мозаичными фрагментами. Это может происходить, если взаимная корреляция между исходным i и целевым j , представленная как $T_x[i][j]$, является очень

низкой для всех i, j .
Например, если:

$$T_x[i][j] < 0,6,$$

причем здесь используется ориентировочное пороговое значение, то:

$$\text{tileNum}[n\text{Tar}]k = \text{tileNum}[n\text{Tar}]k-1$$

для всех $n\text{Tar}$ этого кадра k .

Две вышеуказанные технологии значительно уменьшают артефакты, которые возникают в результате быстрого изменения заданных номеров мозаичных фрагментов между кадрами. Еще одно дополнительное преимущество этого отсечения и стабилизации мозаичных фрагментов состоит в том, что дополнительная информация не должна отправляться в декодер, и при этом не требуется изменение архитектуры декодера. Это предложенное отсечение мозаичных фрагментов является разумным способом уменьшения потенциальных музыкальных шумоподобных артефактов или избыточного шума в мозаичных спектральных областях.

Фиг. 11a иллюстрирует аудиодекодер для декодирования кодированного аудиосигнала. Аудиодекодер содержит (базовый) аудиодекодер 1102 для формирования первого декодированного представления первого набора первых спектральных частей, причем декодированное представление имеет первое спектральное разрешение.

Кроме того, аудиодекодер содержит параметрический декодер 1104 для формирования второго декодированного представления второго набора вторых спектральных частей, имеющих второе спектральное разрешение ниже первого спектрального разрешения. Кроме того, предоставляется модуль 1106 повторного формирования частоты, который принимает, в качестве первого ввода 1101, декодированные первые спектральные части и, в качестве второго ввода в 1103, параметрическую информацию, включающую в себя, для каждого целевого частотного мозаичного фрагмента или целевой полосы частот восстановления, информацию исходного диапазона. Модуль 1106 повторного формирования частоты затем применяет повторное формирование частоты посредством использования спектральных значений из исходного диапазона, идентифицированного посредством информации совпадения, чтобы формировать спектральные данные для целевого диапазона. Затем первые спектральные части 1101 и вывод модуля 1107 повторного формирования частоты вводятся в спектрально-временной преобразователь 1108, чтобы, в завершение, формировать декодированный аудиосигнал.

Предпочтительно, аудиодекодер 1102 представляет собой аудиодекодер в спектральной области, хотя аудиодекодер также может быть реализован как любой другой аудиодекодер, такой как аудиодекодер во временной области или параметрический аудиодекодер.

Как указано на фиг. 11b, модуль 1106 повторного формирования частоты может содержать функциональности блока 1120, иллюстрирующего модуль выбора исходных диапазонов/модулятор мозаичных фрагментов для нечетных запаздываний, отбеливающего фильтра 1122, когда предоставляется флаг 1123 отбеливания, и дополнительно, спектральной огибающей с функциональностями регулирования, реализованными так, как проиллюстрировано в блоке 1128, с использованием необработанных спектральных данных, сформированных либо посредством блока

1120, либо посредством блока 1122, либо посредством взаимодействия обоих блоков. В любом случае, модуль 1106 повторного формирования частоты может содержать переключатель 1124, активируемый в зависимости от принимаемого флага 1123 отбеливания. Когда флаг отбеливания задается, вывод модуля выбора исходных диапазонов/модулятора мозаичных фрагментов для нечетных запаздываний вводится в отбеливающий фильтр 1122. Тем не менее, затем флаг 1123 отбеливания не задается для определенной полосы частот восстановления, в таком случае активируется обходная линия 1126, так что вывод блока 1120 предоставляется в блок 1128 регулирования спектральной огибающей без отбеливания.

Может быть предусмотрено несколько уровней отбеливания (1123), передаваемых в служебных сигналах в потоке битов, и эти уровни могут передаваться в служебных сигналах в расчете на мозаичный фрагмент. В случае если три уровня передаются в служебных сигналах в расчете на мозаичный фрагмент, они должны быть кодированы следующим образом:

```

15  bit=readBit(1);
    if(bit==1) {
        for(tile_index=0...nT)
            /*уровни, идентичные уровню последнего кадра*/
            whitening_level[tile_index]=whitening_level_prev_frame[tile_index];
20  } else {
            /*первый мозаичный фрагмент:*/
            tile_index=0;
            bit=readBit(1);
            if(bit==1) {
25  whitening_level[tile_index]=MID_WHITENING;
            } else {
                bit=readBit(1);
                if(bit==1) {
30  whitening_level[tile_index]=STRONG_WHITENING;
                } else {
                    whitening_level[tile_index]=OFF; /*без отбеливания*/
                }
            }
            /*оставшиеся мозаичные фрагменты:*/
35  bit=readBit(1);
            if(bit==1) {
                /*уровни сглаживания для оставшихся мозаичных фрагментов являются идентичным
                уровню для первого*/
                /*дополнительные биты не должны считываться*/
40  for(tile_index=1...nT)
                    whitening_level[tile_index]=whitening_level[0];
            } else {
                /*считывание битов для оставшихся мозаичных фрагментов, как для первого
                мозаичного фрагмента*/
45  for(tile_index=1...nT) {
                    bit=readBit(1);
                    if(bit==1) {
                        whitening_level[tile_index]=MID_WHITENING;

```



```

    } else {
    bit=readBit(1);
    if(bit==1) {
    whitening_level[tile_index]=STRONG_WHITENING;
5    } else {
    whitening_level[tile_index]=OFF;/*без отбеливания*/
    }
    }
    }
10    }
    }

```

MID_WHITENING и STRONG_WHITENING означают различные отбеливающие фильтры (1122), которые могут отличаться по способу, которым вычисляется огибающая (как описано выше).

15 Модуль повторного формирования частоты на стороне декодера может управляться посредством идентификатора 1121 исходного диапазона, когда применяется только схема приблизительного выбора спектральных мозаичных фрагментов. Тем не менее, когда применяется схема настраиваемого выбора спектральных мозаичных фрагментов, в таком случае, дополнительно, предоставляется запаздывание 1119 для исходного

20 диапазона. Кроме того, при условии, что вычисление корреляции предоставляет отрицательный результат, в таком случае, дополнительно, знак корреляции также может применяться к блоку 1120, так что спектральные линии страничных данных умножаются на -1, с тем чтобы учитывать знак минус.

Таким образом, настоящее изобретение, как пояснено на фиг. 11a, 11b, удостоверяется,

25 что получено оптимальное качество звука вследствие того факта, что наилучшее совпадающий исходный диапазон для определенного намеченного или целевого диапазона вычисляется на стороне кодера и применяется на стороне декодера.

Фиг. 11c является определенным аудиокодером для кодирования аудиосигнала, содержащим временно-спектральный преобразователь 1130, последующий соединенный

30 спектральный анализатор 1132 и, дополнительно, модуль 1134 вычисления параметров и базовый кодер 1136. Базовый кодер 1136 выводит кодированные исходные диапазоны, и модуль 1134 вычисления параметров выводит информацию совпадения для целевых диапазонов.

Кодированные исходные диапазоны передаются в декодер вместе с информацией

35 совпадения для целевых диапазонов, так что декодер, проиллюстрированный на фиг. 11a, находится в позиции для того, чтобы выполнять повторное формирование частоты.

Модуль 1134 вычисления параметров выполнен с возможностью вычисления подобий между первыми спектральными частями и вторыми спектральными частями и с

возможностью определения, на основе вычисленных подобий для второй спектральной

40 части, совпадающей первой спектральной части, совпадающей со второй спектральной частью. Предпочтительно, совпадающие результаты для различных исходных диапазонов и целевых диапазонов, как проиллюстрировано на фиг. 12a, 12b, чтобы определять выбранную совпадающую пару, содержащую вторую спектральную часть, и модуль вычисления параметров выполнен с возможностью предоставления этой

45 информации совпадения, идентифицирующей совпадающую пару, в кодированный аудиосигнал. Предпочтительно, этот модуль 1134 вычисления параметров выполнен с возможностью использования предварительно заданных целевых областей во втором наборе вторых спектральных частей или предварительно заданных исходных областей

в первом наборе первых спектральных частей, как проиллюстрировано, например, на фиг. 12b. Предпочтительно, предварительно заданные целевые области являются
 5 неперекрывающимися, или предварительно заданные исходные области являются перекрывающимися. Когда предварительно заданные исходные области представляют собой поднабор первого набора первых спектральных частей ниже начальной частоты
 309 заполнения интервалов по фиг. 3a, и предпочтительно, предварительно заданная целевая область, покрывающая нижнюю спектральную область, совпадает, в нижней частотной границе, с начальной частотой заполнения интервалов, так что все целевые диапазоны расположены выше начальной частоты заполнения интервалов, а исходные
 10 диапазоны расположены ниже начальной частоты заполнения интервалов.

Как пояснено, сверхвысокая степень детализации получена посредством сравнения целевой области с исходной областью без запаздывания для исходной области и с идентичной исходной областью, но с определенным запаздыванием. Эти запаздывания применяются в модуле 1140 вычисления взаимной корреляции по фиг. 11d, и выбор
 15 совпадающей пары, в завершение, выполняется посредством модуля 1144 выбора мозаичных фрагментов.

Кроме того, предпочтительно выполнять отбеливание исходных и/или целевых диапазонов, проиллюстрированное в блоке 1142. Этот блок 1142 затем предоставляет
 20 флаг отбеливания для потока битов, который используется для управления переключателем 1123 на стороне декодера по фиг. 11b. Кроме того, если модуль 1140 вычисления взаимной корреляции предоставляет отрицательный результат, то этот отрицательный результат также передается в служебных сигналах в декодер. Таким образом, в предпочтительном варианте осуществления, модуль выбора мозаичных фрагментов выводит идентификатор исходного диапазона для целевого диапазона,
 25 запаздывание, знак, а блок 1142 дополнительно предоставляет флаг отбеливания.

Кроме того, модуль 1134 вычисления параметров выполнен с возможностью осуществления отсечения 1146 исходных мозаичных фрагментов посредством уменьшения числа потенциальных исходных диапазонов так, что исходное наложение
 30 отбрасывается из набора потенциальных исходных мозаичных фрагментов на основе порогового значения подобия. Таким образом, когда два исходных мозаичных фрагмента являются подобными со значением, большим или равным пороговому значению подобия, то один из двух исходных мозаичных фрагментов удаляется из набора потенциальных источников, и удаленный исходный мозаичный фрагмент более
 35 не используется для последующей обработки и, в частности, не может выбираться посредством модуля 1144 выбора мозаичных фрагментов, либо не используется для вычисления взаимной корреляции между различными исходными диапазонами и целевыми диапазонами, выполняемого в блоке 1140.

Различные реализации описаны относительно различных чертежей. Фиг. 1a-5c связаны со схемой полноскоростного (с полной полосой пропускания) кодера/декодера. Фиг.
 40 6a-7e связаны со схемой кодера/декодера с TNS- или TTS-обработкой. Фиг. 8a-8e связаны со схемой кодера/декодера с конкретной двухканальной обработкой. Фиг. 9a-10d связаны с конкретным вычислением и применением информации энергии, а фиг. 11a-12c связаны с конкретным способом выбора мозаичных фрагментов.

Все эти различные аспекты могут иметь изобретательское применение независимое друг от друга, но, дополнительно, также могут применяться совместно, как, по существу,
 45 проиллюстрировано на фиг. 2a и 2b. Тем не менее, конкретная двухканальная обработка также может применяться к схеме кодера/декодера, проиллюстрированной на фиг. 13, и то же является истинным для TNS/TTS-обработки, вычисления и применения

информации энергии огибающей в полосе частот восстановления или адаптивной идентификации исходного диапазона и соответствующего применения на стороне декодера. С другой стороны, полноскоростной аспект может применяться с/без TNS/TTS-обработки, с/без двухканальной обработки, с/без адаптивной идентификации исходного диапазона либо с другими видами вычислений энергии для представления спектральной огибающей. Таким образом, очевидно, что признаки одного из этих отдельных аспектов также могут применяться в других аспектах.

Хотя некоторые аспекты описаны в контексте устройства для кодирования или декодирования, очевидно, что эти аспекты также представляют описание соответствующего способа, при этом блок или устройство соответствует этапу способа либо признаку этапа способа. Аналогично, аспекты, описанные в контексте этапа способа, также представляют описание соответствующего блока или элемента, или признака соответствующего устройства. Некоторые или все этапы способа могут быть выполнены посредством (или с использованием) устройства, такого как, например, микропроцессор, программируемый компьютер либо электронная схема. В некоторых вариантах осуществления, некоторые из одного или более самых важных этапов способа могут выполняться посредством этого устройства.

В зависимости от определенных требований к реализации, варианты осуществления изобретения могут быть реализованы в аппаратных средствах или в программном обеспечении. Реализация может выполняться с использованием энергонезависимого носителя хранения данных, такого как цифровой носитель хранения данных, например, гибкий диск, жесткий диск (HDD), DVD, Blu-Ray, CD, ROM, PROM и EPROM, EEPROM или флэш-память, имеющего сохраненные электронночитаемые управляющие сигналы, которые взаимодействуют (или допускают взаимодействие) с программируемой компьютерной системой таким образом, что осуществляется соответствующий способ. Следовательно, цифровой носитель хранения данных может быть машиночитаемым.

Некоторые варианты осуществления согласно изобретению содержат носитель данных, имеющий электронночитаемые управляющие сигналы, которые допускают взаимодействие с программируемой компьютерной системой таким образом, что осуществляется один из способов, описанных в данном документе.

В общем, варианты осуществления настоящего изобретения могут быть реализованы как компьютерный программный продукт с программным кодом, при этом программный код выполнен с возможностью осуществления одного из способов, когда компьютерный программный продукт работает на компьютере. Программный код, например, может быть сохранен на машиночитаемом носителе.

Другие варианты осуществления содержат компьютерную программу для осуществления одного из способов, описанных в данном документе, сохраненную на машиночитаемом носителе.

Другими словами, следовательно, вариант осуществления изобретаемого способа представляет собой компьютерную программу, имеющую программный код для осуществления одного из способов, описанных в данном документе, когда компьютерная программа работает на компьютере.

Следовательно, дополнительный вариант осуществления изобретаемого способа представляет собой носитель хранения данных (цифровой носитель хранения данных или машиночитаемый носитель), содержащий записанную компьютерную программу для осуществления одного из способов, описанных в данном документе. Носитель данных, цифровой носитель хранения данных или носитель с записанными данными типично является материальным и/или энергонезависимым.

Следовательно, дополнительный вариант осуществления изобретаемого способа представляет собой поток данных или последовательность сигналов, представляющих компьютерную программу для осуществления одного из способов, описанных в данном документе. Поток данных или последовательность сигналов, например, может быть
5 выполнена с возможностью передачи через соединение для передачи данных, например, через Интернет.

Дополнительный вариант осуществления содержит средство обработки, например, компьютер или программируемое логическое устройство, выполненное с возможностью осуществлять один из способов, описанных в данном документе.

10 Дополнительный вариант осуществления содержит компьютер, имеющий установленную компьютерную программу для осуществления одного из способов, описанных в данном документе.

Дополнительный вариант осуществления согласно изобретению содержит устройство или систему, выполненную с возможностью передавать (например, электронно или
15 оптически) компьютерную программу для осуществления одного из способов, описанных в данном документе, в приемное устройство. Приемное устройство, например, может представлять собой компьютер, мобильное устройство, запоминающее устройство и т.п. Устройство или система, например, может содержать файловый сервер для передачи компьютерной программы в приемное устройство.

20 В некоторых вариантах осуществления, программируемое логическое устройство (например, программируемая пользователем вентильная матрица) может быть использовано для того, чтобы выполнять часть или все из функциональностей способов, описанных в данном документе. В некоторых вариантах осуществления, программируемая пользователем вентильная матрица может взаимодействовать с
25 микропроцессором, чтобы осуществлять один из способов, описанных в данном документе. В общем, способы предпочтительно осуществляются посредством любого устройства.

Вышеописанные варианты осуществления являются просто иллюстративными в отношении принципов настоящего изобретения. Следует понимать, что модификации
30 и изменения компоновок и подробностей, описанных в данном документе, должны быть очевидными для специалистов в данной области техники. Следовательно, они подразумеваются как ограниченные только посредством объема нижеприведенной формулы изобретения, а не посредством конкретных подробностей, представленных посредством описания и пояснения вариантов осуществления в данном документе.

35 Список библиографических ссылок

[1] Dietz, L. Liljeryd, K. Kjörling и O. Kunz "Spectral Band Replication, the novel approach in audio coding", in 112th AES Convention, Мюнхен, май 2002 года.

[2] Ferreira, D. Sinha "Accurate Spectral Replacement", Audio Engineering Society Convention, Барселона, Испания, 2005 год.

40 [3] D. Sinha, A. Ferreira и E. Harinarayanan "A Novel Integrated Audio Bandwidth Extension Toolkit (ABET)", Audio Engineering Society Convention, Париж, Франция, 2006 год.

[4] R. Annadana, E. Harinarayanan, A. Ferreira и D. Sinha "New Results in Low Bit Rate Speech Coding и Bandwidth Extension", Audio Engineering Society Convention, Сан-Франциско, США, 2006 год.

45 [5] T. Żernicki, M. Bartkowiak "Audio bandwidth extension by frequency scaling of sinusoidal partials", Audio Engineering Society Convention, Сан-Франциско, США, 2008 год.

[6] J. Herre, D. Schulz "Extending the MPEG-4 AAC Codec by Perceptual Noise Substitution", 104th AES Convention, Амстердам, 1998 год, Preprint 4720.

[7] M. Neuendorf, M. Multrus, N. Rettelbach и др. "MPEG Unified Speech и Audio Coding-The ISO/MPEG Standard for High-Efficiency Audio Coding of all Content Types", 132nd AES Convention, Будапешт, Венгрия, апрель 2012 года.

[8] McAulay, Robert J., Quatieri, Thomas F. "Speech Analysis/Synthesis Based on the Sinusoidal Representation". IEEE Transactions on Acoustics, Speech и Signal Processing, издание 34(4), август 1986 года.

[9] Smith, J.O., Serra, X "PARSHL: An analysis/synthesis program for non-harmonic sounds based on the sinusoidal representation", Proceedings of the International Computer Music Conference, 1987 год.

[10] Purnhagen, H.; Meine, Nikolaus "HILN-the MPEG-4 parametric audio coding tools", Circuits и Systems, 2000 Proceedings ISCAS, 2000 год, Женева, The 2000 IEEE International Symposium on, издание 3, номер, стр. 201, 204, том 3, 2000 год.

[11] International Standard ISO/IEC 13818-3 "Generic Coding of Moving Pictures и Associated Audio: Audio", Женева, 1998 год.

[12] M. Bosi, K. Brandenburg, S. Quackenbush, L. Fielder, K. Akagiri, H. Fuchs, M. Dietz, J. Herre, G. Davidson, Oikawa "MPEG-2 Advanced Audio Coding", 101st AES Convention, Лос-Анжелес, 1996 год

[13] J. Herre "Temporal Noise Shaping, Quantization и Coding methods in Perceptual Audio Coding: A Tutorial introduction", 17th AES International Conference on High Quality Audio Coding, август 1999 года.

[14] J. Herre "Temporal Noise Shaping, Quantization и Coding methods in Perceptual Audio Coding: A Tutorial introduction", 17th AES International Conference on High Quality Audio Coding, август 1999 года.

[15] International Standard ISO/IEC 23001-3:2010 "Unified speech и audio coding Audio", Женева, 2010 год.

[16] International Standard ISO/IEC 14496-3:2005 "Information technology – Coding of audio-visual objects – Part 3: Audio", Женева, 2005 год.

[17] P. Ekstrand "Bandwidth Extension of Audio Signals by Spectral Band Replication", in Proceedings of 1st IEEE Benelux Workshop on MPCA, Leuven, ноябрь 2002 года.

[18] F. Nagel, S. Disch, S. Wilde "A continuous modulated single sideband bandwidth extension", ICASSP International Conference on Acoustics, Speech и Signal Processing, Даллас, Texas (США), апрель 2010 года.

(57) Формула изобретения

1. Устройство для формирования декодированного двухканального сигнала, содержащее:

- аудиопроцессор (802) для декодирования кодированного двухканального сигнала для того, чтобы получать первый набор первых спектральных частей;

- параметрический декодер (804) для предоставления параметрических данных для второго набора вторых спектральных частей и двухканального идентификатора, идентифицирующего первое двухканальное представление для второй спектральной части из второго набора вторых спектральных частей или второе другое двухканальное представление для второй спектральной части из второго набора вторых спектральных частей; и

- модуль (806) повторного формирования частоты для повторного формирования второй спектральной части из второго набора вторых спектральных частей в зависимости от первой спектральной части из первого набора первых спектральных частей, параметрических данных для второй спектральной части из второго набора

вторых спектральных частей и двухканального идентификатора для второй спектральной части из второго набора вторых спектральных частей,

- при этом аудиопроцессор выполнен с возможностью декодировать первый набор первых спектральных частей в соответствии с дополнительным двухканальным идентификатором для первого набора первых спектральных частей, причем дополнительный двухканальный идентификатор идентифицирует первое двухканальное представление для первой спектральной части из первого набора первых спектральных частей или второе другое двухканальное представление для первой спектральной части из первого набора первых спектральных частей, и преобразовывать (832) первый набор первых спектральных частей таким образом, что получаются первое двухканальное представление первых спектральных частей из первого набора первых спектральных частей и второе двухканальное представление первых спектральных частей из первого набора первых спектральных частей, и

- при этом модуль (806) повторного формирования частоты выполнен с возможностью использовать либо первое двухканальное представление первого набора первых спектральных частей, либо второе двухканальное представление первого набора первых спектральных частей, как указано в двухканальном идентификаторе для второй спектральной части из второго набора вторых спектральных частей.

2. Устройство по п. 1, в котором двухканальный идентификатор идентифицирует либо отдельную обработку двух каналов кодированного двухканального сигнала, либо объединенную обработку двух каналов кодированного двухканального сигнала, и

- при этом модуль (806) повторного формирования частоты выполнен с возможностью повторного формирования второй спектральной части из второго набора вторых спектральных частей для первого канала из двух каналов и второй спектральной части из второго набора вторых спектральных частей для второго канала из двух каналов с использованием первой части первого канала и первой части второго канала, при этом первая часть первого канала и первая часть второго канала находятся в двухканальном представлении, идентифицированном посредством двухканального идентификатора для второй спектральной части из второго набора вторых спектральных частей.

3. Устройство по п. 1,

- в котором двухканальный идентификатор идентифицирует либо отдельную обработку двух каналов кодированного двухканального сигнала, либо объединенную обработку двух каналов кодированного двухканального сигнала, и

- при этом модуль (806) повторного формирования частоты выполнен с возможностью повторного формирования объединенного представления двух каналов во второй спектральной части из второго набора вторых спектральных частей, идентифицированного посредством двухканального идентификатора, и

- при этом модуль (806) повторного формирования частоты дополнительно содержит модуль (842) преобразования представлений для преобразования объединенного представления второй спектральной части из второго набора вторых спектральных частей в отдельное представление для второй спектральной части из второго набора вторых спектральных частей.

4. Устройство по п. 3, в котором модуль преобразования представлений использует дополнительные параметры объединенного представления для преобразования представления.

5. Устройство по п. 3, в котором объединенное представление является средним/боковым представлением, при этом модуль преобразования представлений выполнен

с возможностью работать на основе следующего уравнения:

$$leftTile = 0.5 * (midTile + sideTile)$$

$$rightTile = 0.5 * (midTile - sideTile)$$

- где leftTile и rightTile являются отдельным представлением для второй спектральной части из второго набора вторых спектральных частей и где midTile и sideTile являются объединенным представлением для второй спектральной части из второго набора вторых спектральных частей.

6. Устройство по п. 3, в котором объединенное представление является представлением, содержащим дополнительный коэффициент прогнозирования, при этом модуль преобразования представлений выполнен с возможностью работать на основе следующего уравнения:

$$sideTile[k] = sideTile[k] - predictionCoeff \cdot midTile[k]$$

$$leftTile[k] = midTile[k] + sideTile[k]$$

$$rightTile[k] = midTile[k] - sideTile[k]$$

- когда направление прогнозирования задается от среднего к боковому, или

$$midTile1[k] = midTile[k] - predictionCoeff \cdot sideTile[k]$$

$$leftTile[k] = midTile1[k] - sideTile[k]$$

$$rightTile[k] = midTile1[k] + sideTile[k]$$

- когда направление прогнозирования указывается от бокового к среднему,

- где leftTile и rightTile являются отдельным представлением для второй спектральной части из второго набора вторых спектральных частей, и где midTile и sideTile являются объединенным представлением для второй спектральной части из второго набора вторых спектральных частей, и где predictionCoefficient является дополнительным коэффициентом прогнозирования.

7. Устройство по п. 1,

- в котором параметрические данные для второго набора вторых спектральных частей отдельно приводятся для каждого канала двухканального представления, и

- при этом модуль (806) повторного формирования частоты выполнен с возможностью преобразования (840) параметрических данных для второй спектральной части из второго набора вторых спектральных частей в объединенное представление для второй спектральной части из второго набора вторых спектральных частей и применения параметрических данных к объединенному представлению первой спектральной части из первого набора первых спектральных частей, когда двухканальный идентификатор идентифицирует объединенное представление для второй спектральной части из второго набора вторых спектральных частей.

8. Устройство по п. 1, в котором вторые спектральные части из второго набора вторых спектральных частей соответствуют полосам частот, при этом двухканальный идентификатор представляет собой массив флагов, по одному флагу для каждой полосы частот, при этом параметрический декодер (804) выполнен с возможностью проверять то, задается флаг или нет, и управлять повторным формированием частоты в соответствии с флагом, чтобы использовать либо первое двухканальное представление, либо второе двухканальное представление первой спектральной части из первого набора первых спектральных частей кодированного двухканального сигнала.

9. Устройство по п. 1, в котором параметрический декодер (804) выполнен с возможностью предоставлять дополнительный двухканальный идентификатор для

первого набора первых спектральных частей, указывающих первое или второе другое двухканальное представление для первой спектральной части из первого набора первых спектральных частей, и

- при этом устройство выполнено с возможностью декодирования второго

5 двухканального представления, как указано посредством двухканального идентификатора для первой спектральной части из первого набора первых спектральных частей, и

- при этом модуль (806) повторного формирования частоты выполнен с возможностью преобразования (832) второго двухканального представления в первое
10 двухканальное представление после базового декодирования (830).

10. Устройство по п. 1, дополнительно содержащее модуль (846) комбинирования для комбинирования первого набора первых спектральных частей, сформированных посредством аудиопроцессора (802), и восстановленной второй спектральной части из второго набора вторых спектральных частей, сформированной посредством модуля
15 (806) повторного формирования частоты, чтобы получать декодированный двухканальный сигнал.

11. Устройство по п. 1,

- в котором параметрический декодер выполнен с возможностью дополнительно предоставлять для второй спектральной части из второго набора вторых спектральных
20 частей идентификатор (1121) исходной полосы частот, указывающий конкретную первую спектральную часть из первого набора первых спектральных частей, которая должна быть использована для повторного формирования второй спектральной части из второго набора вторых спектральных частей, и

- при этом модуль (806) повторного формирования частоты выполнен с
25 возможностью повторно формировать вторую спектральную часть из второго набора вторых спектральных частей с использованием первой спектральной части из первого набора первых спектральных частей, идентифицированной посредством идентификатора (833) исходной полосы частот.

12. Устройство по п. 1,

- при этом модуль (806) повторного формирования частоты содержит модуль (832) преобразования представлений для предоставления первого и второго двухканального
30 представления первого набора первых спектральных частей, сформированных посредством аудиопроцессора (802),

- при этом модуль (806) повторного формирования частоты дополнительно содержит
35 модуль (836) формирования частотных мозаичных фрагментов для формирования необработанных данных для каждого канала канального представления, идентифицированного посредством двухканального идентификатора, и использования идентификатора (833) исходного диапазона, указывающего первые спектральные части из первого набора первых спектральных частей, которые должны быть использованы
40 для формирования необработанных данных,

- при этом модуль (806) повторного формирования частоты дополнительно содержит модуль (840) преобразования параметров для преобразования параметров, предоставленных в первом двухканальном представлении, во второе двухканальное
представление для параметров, когда необработанные данные для каждого канала
45 предоставляются во втором двухканальном представлении посредством модуля (836) формирования частотных мозаичных фрагментов,

- при этом модуль (806) повторного формирования частоты дополнительно содержит модуль (838) регулирования огибающей для регулирования огибающей каждого канала

двухканального представления, причем двухканальное представление представляет собой второе двухканальное представление,

- при этом модуль (806) повторного формирования частоты дополнительно содержит модуль (842) преобразования представлений для преобразования двухканального представления спектральных значений во второй спектральной части из второго набора вторых спектральных частей в первое двухканальное представление,

- при этом устройство дополнительно содержит частотно-временной преобразователь (846) для преобразования представления, сформированного посредством модуля (842) преобразования представлений, из спектральной области во временную область.

13. Аудиодекодер по п. 1, в котором первое двухканальное представление для второй спектральной части из второго набора вторых спектральных частей и второе другое двухканальное представление для второй спектральной части из второго набора вторых спектральных частей выбираются из группы двухканальных представлений, содержащей левое/правое двухканальное представление, среднее/боковое двухканальное представление и двухканальное представление низведение/остаток, и

при этом первое двухканальное представление для первой спектральной части из первого набора первых спектральных частей и второе другое двухканальное представление для первой спектральной части из первого набора первых спектральных частей выбираются из группы двухканальных представлений, содержащей левое/правое двухканальное представление, среднее/боковое двухканальное представление и двухканальное представление низведение/остаток.

14. Аудиокодер для кодирования двухканального аудиосигнала, содержащий:

- временно-спектральный преобразователь (860) для преобразования двухканального аудиосигнала в спектральное представление;

- спектральный анализатор (866) для предоставления индикатора первого набора первых спектральных частей, которые должны быть кодированы с первым спектральным разрешением, и второго набора вторых спектральных частей, которые должны быть кодированы посредством второго спектрального разрешения, причем второе спектральное разрешение меньше первого спектрального разрешения,

- двухканальный анализатор (864) для анализа полос частот двухканального аудиосигнала в диапазоне восстановления, чтобы определять двухканальный идентификатор для полосы частот восстановления, причем двухканальный идентификатор идентифицирует первое двухканальное представление для второй спектральной части из второго набора вторых спектральных частей в полосе частот восстановления или второе другое двухканальное представление для второй спектральной части из второго набора вторых спектральных частей в полосе частот восстановления, при этом полоса частот восстановления содержит вторую спектральную часть из второго набора вторых спектральных частей;

- базовый кодер (870) для кодирования первого набора первых спектральных частей, чтобы предоставлять первое кодированное представление; и

- модуль (868) кодирования параметров для вычисления параметрических данных в отношении второго набора вторых спектральных частей, чтобы получать кодированное параметрическое представление,

- при этом кодированный двухканальный аудиосигнал содержит первое кодированное представление, кодированное параметрическое представление и двухканальный идентификатор для второй спектральной части из второго набора вторых спектральных частей в полосе частот восстановления.

15. Аудиокодер по п. 14, дополнительно содержащий модуль (862) преобразования

для каждой полосы частот для преобразования первых спектральных частей из первого набора первых спектральных частей в двухканальное представление, указываемое посредством двухканального идентификатора, при этом спектральный анализатор (866) выполнен с возможностью анализа двухканального представления, выводимого посредством модуля (862) преобразования для каждой полосы частот.

16. Аудиокодер по п. 14,

- в котором двухканальный анализатор (864) выполнен с возможностью выполнения вычисления корреляции между второй спектральной частью из второго набора вторых спектральных частей первого канала двухканального представления и второй спектральной частью из второго набора вторых спектральных частей двух каналов двухканального представления, чтобы определять либо отдельное двухканальное представление, либо объединенное двухканальное представление.

17. Аудиокодер по п. 14,

- в котором спектральный анализатор (866) выполнен с возможностью сравнения совпадающих результатов для различных спектральных частей по меньшей мере одного канала двухканального представления с другой первой спектральной частью из первого набора первых спектральных частей по меньшей мере одного канала двухканального представления, чтобы определять совпадающую пару из первой спектральной части из первого набора первых спектральных частей по меньшей мере одного канала и второй спектральной части из второго набора вторых спектральных частей по меньшей мере одного канала и предоставлять индикатор (833) совпадения для наилучшей совпадающей пары, и

- при этом аудиокодер выполнен с возможностью выводить, в дополнение к кодированному двухканальному аудиосигналу, индикатор (833) совпадения для второй спектральной части из второго набора вторых спектральных частей.

18. Аудиокодер по п. 14, дополнительно содержащий модуль (862) преобразования для каждой полосы частот, имеющий вход, соединенный с выходом временно-спектрального преобразователя (860),

- в котором спектральный анализатор (866) выполнен с возможностью принимать в качестве ввода вывод модуля (862) преобразования для каждой полосы частот;

- при этом двухканальный анализатор (864) выполнен с возможностью анализа вывода временно-спектрального преобразователя (860) и предоставления результата анализа, чтобы управлять модулем (862) преобразования для каждой полосы частот,

- при этом аудиокодер выполнен с возможностью кодировать вывод модуля (862) преобразования для каждой полосы частот, управляемого посредством спектрального анализатора (866), так что только первый набор первых спектральных частей из первого набора первых спектральных частей кодируется посредством базового кодера (870), и

- при этом модуль (868) вычисления параметров выполнен с возможностью параметрически кодировать второй набор вторых спектральных частей, как указано посредством спектрального анализатора (866) в выводе модуля (862) преобразования для каждой полосы частот.

19. Способ формирования декодированного двухканального сигнала, содержащий этапы, на которых:

- декодируют (802) кодированный двухканальный сигнал, чтобы получать первый набор первых спектральных частей;

- предоставляют (804) параметрические данные для второго набора вторых спектральных частей и двухканальный идентификатор, идентифицирующий первое

двухканальное представление для вторых спектральных частей или второе другое двухканальное представление для второй спектральной части из второго набора вторых спектральных частей; и

- повторно формируют (806) вторую спектральную часть из второго набора вторых спектральных частей в зависимости от первой спектральной части из первого набора первых спектральных частей, параметрических данных для второй спектральной части из второго набора вторых спектральных частей и двухканального идентификатора для второй спектральной части из второго набора вторых спектральных частей,

- при этом декодирование (802) содержит этап, на котором декодируют первый набор первых спектральных частей в соответствии с дополнительным двухканальным идентификатором для первого набора первых спектральных частей, причем дополнительный двухканальный идентификатор идентифицирует первое двухканальное представление для первой спектральной части из первого набора первых спектральных частей или второе другое двухканальное представление для первой спектральной части из первого набора первых спектральных частей, и преобразуют (832) первый набор первых спектральных частей таким образом, что получаются первое двухканальное представление первых спектральных частей из первого набора первых спектральных частей и второе двухканальное представление первых спектральных частей из первого набора первых спектральных частей, и

- при этом повторное формирование (806) содержит этап, на котором используют либо первое двухканальное представление первого набора первых спектральных частей, либо второе двухканальное представление первого набора первых спектральных частей, как указано в двухканальном идентификаторе для второй спектральной части из второго набора вторых спектральных частей.

20. Способ кодирования двухканального аудиосигнала, содержащий этапы, на которых:

- преобразуют (860) двухканальный аудиосигнал в спектральное представление;

- предоставляют (866) индикатор первого набора первых спектральных частей,

которые должны быть кодированы с первым спектральным разрешением, и второго набора вторых спектральных частей, которые должны быть кодированы посредством второго спектрального разрешения, причем второе спектральное разрешение меньше первого спектрального разрешения;

- анализируют (864) полосы частот двухканального аудиосигнала в диапазоне восстановления, чтобы определять двухканальный идентификатор для полосы частот восстановления, причем двухканальный идентификатор идентифицирует первое двухканальное представление для второй спектральной части из второго набора вторых спектральных частей в полосе частот восстановления или второе другое двухканальное представление для второй спектральной части из второго набора вторых спектральных частей в полосе частот восстановления, при этом полоса частот восстановления

содержит вторую спектральную часть из второго набора вторых спектральных частей;

- кодируют (870) первый набор первых спектральных частей, чтобы предоставлять первое кодированное представление; и

- вычисляют (868) параметрические данные для второго набора вторых спектральных частей, чтобы получать кодированное параметрическое представление,

- при этом кодированный двухканальный аудиосигнал содержит первое кодированное представление, кодированное параметрическое представление и двухканальный идентификатор для второй спектральной части из второго набора вторых спектральных частей в полосе частот восстановления.

21. Машиночитаемый носитель, хранящий компьютерную программу для осуществления, при выполнении на компьютере или процессоре, способа по п. 19.

22. Машиночитаемый носитель, хранящий компьютерную программу для осуществления, при выполнении на компьютере или процессоре, способа по п. 20.

5

10

15

20

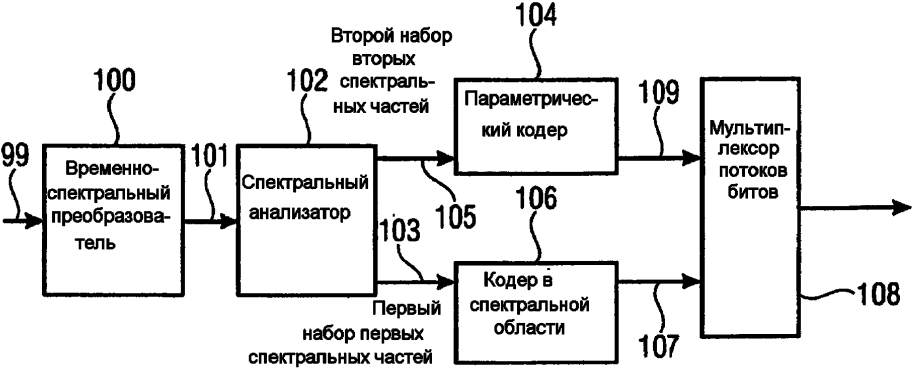
25

30

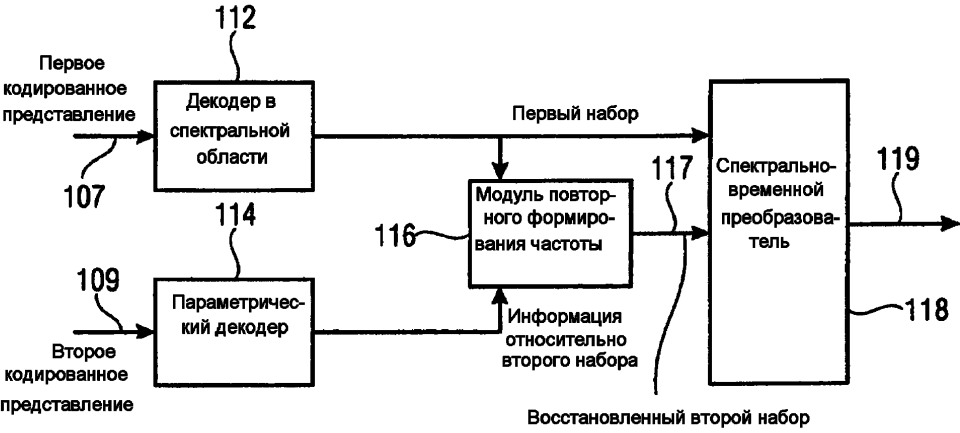
35

40

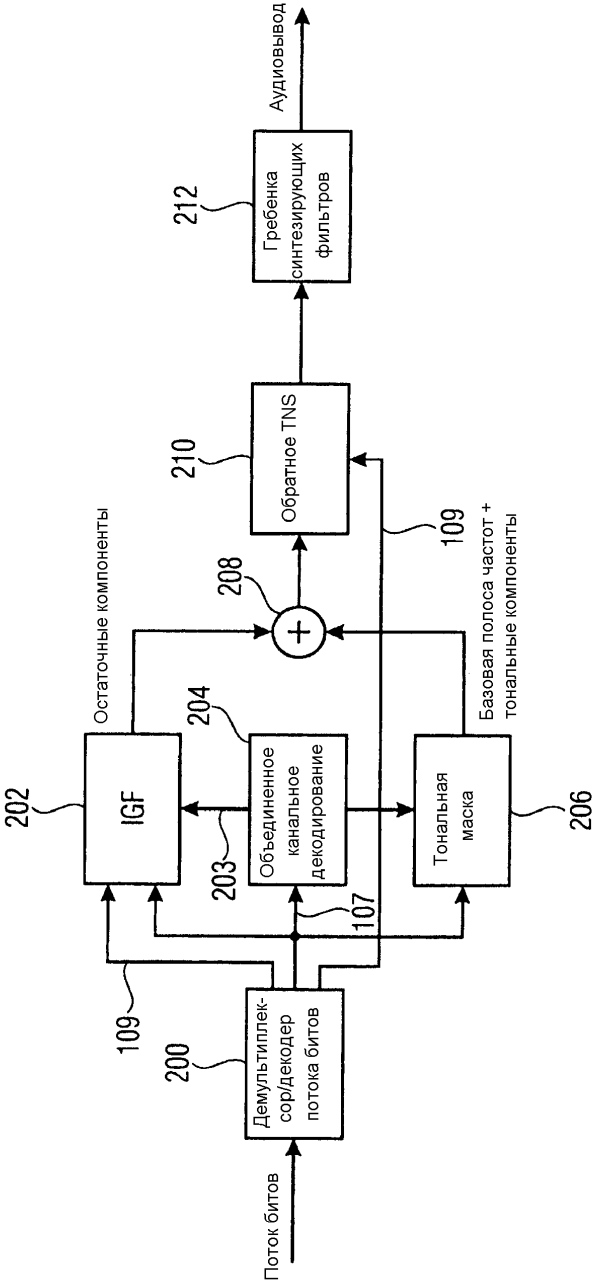
45



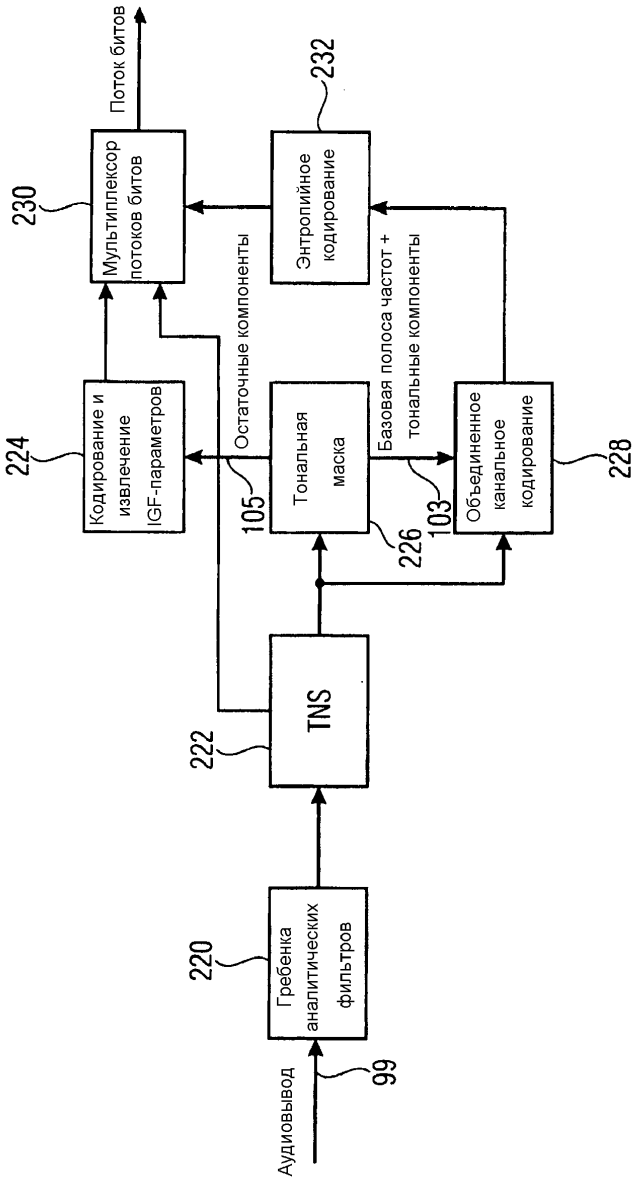
Фиг.1a



Фиг.1b

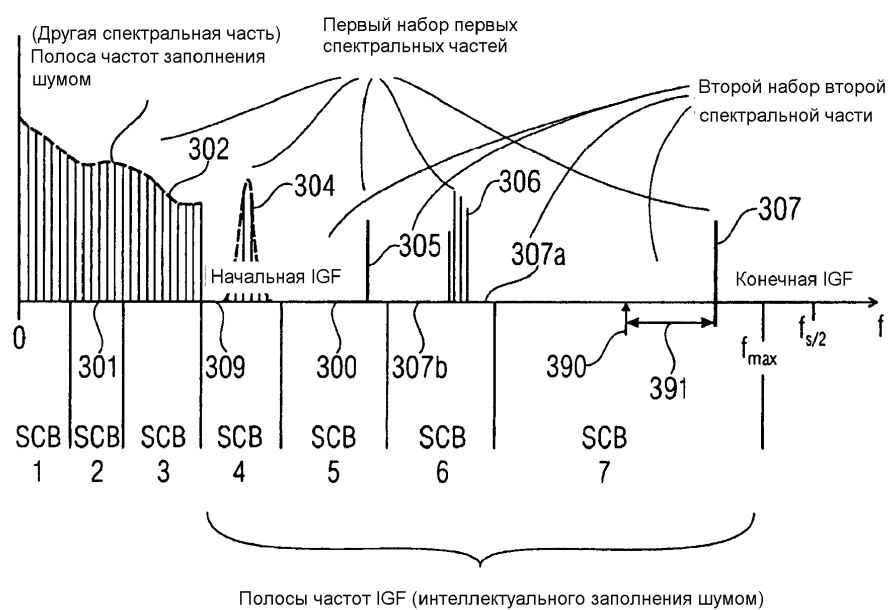


Фиг. 2а



Фиг.2b

- Первое разрешение (высокое разрешение): для "оглающей" первого набора (полинейное кодирование)
- Второе разрешение (низкое разрешение): для "оглающей" второго набора (коэффициент масштабирования в расчете на SCB)



Фиг.3а

| | | | | | | |
|------|-----------------|------|----------------|----------------|----------------|----------------|
| SCB1 | SCB2 | SCB3 | SCB4 | SCB5 | SCB6 | SCB7 |
| SF1 | SF2 | SF3 | SF4 | SF5 | SF6 | SF7 |
| | | | E ₁ | E ₂ | E ₃ | E ₄ |
| | NF ₂ | | | | | |

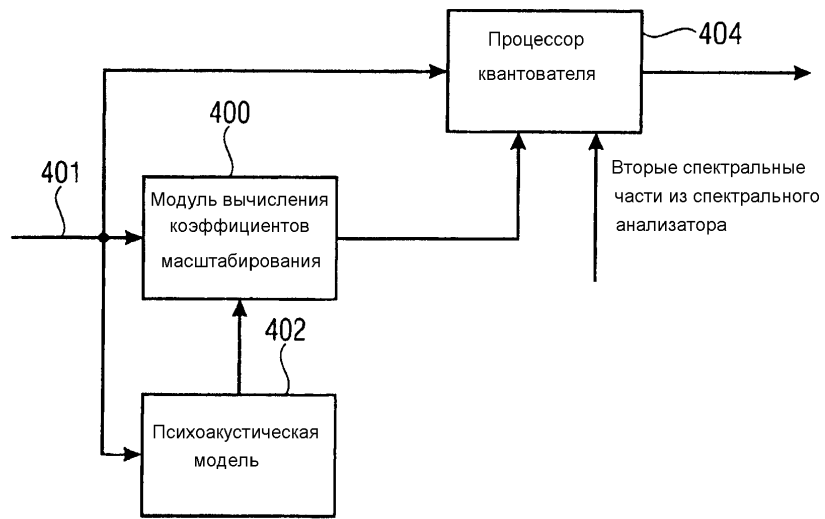
308

310

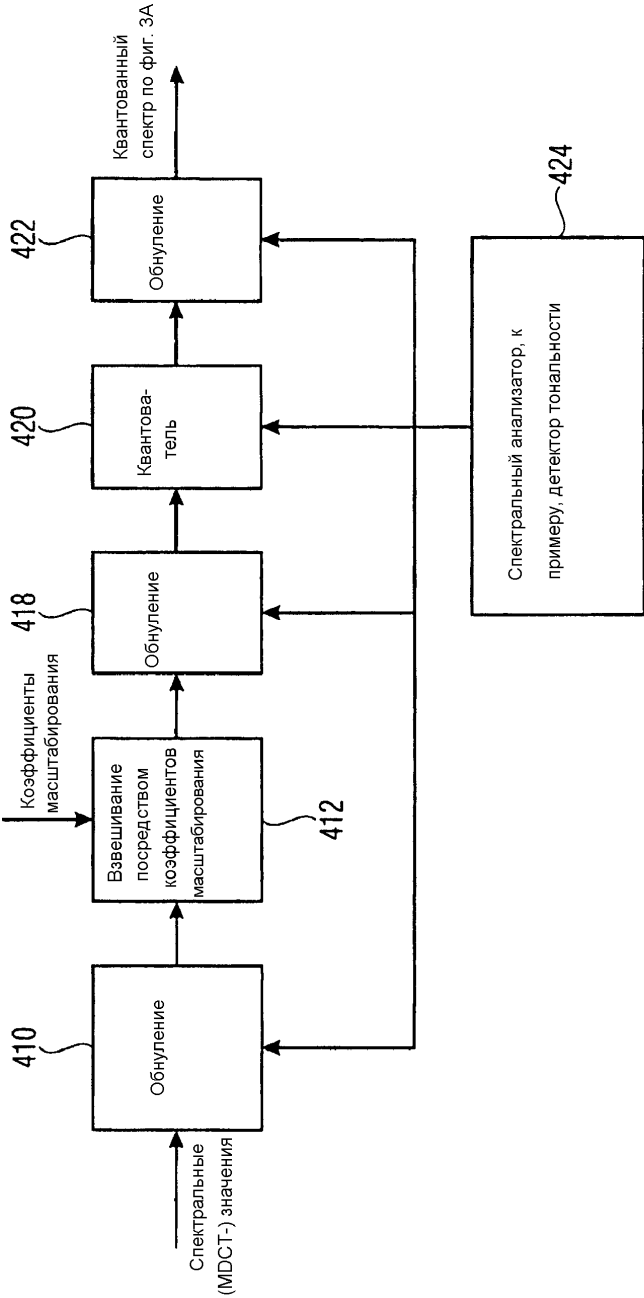
312

Фиг.3b

6/29



Фиг.4а



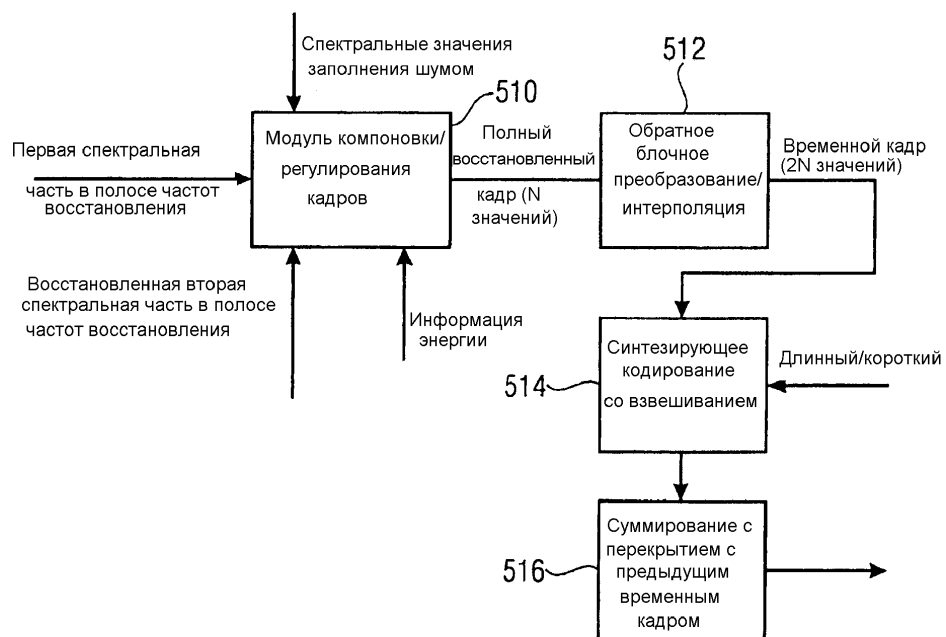
(Процессор квантователя)
Фиг. 4b

8/29

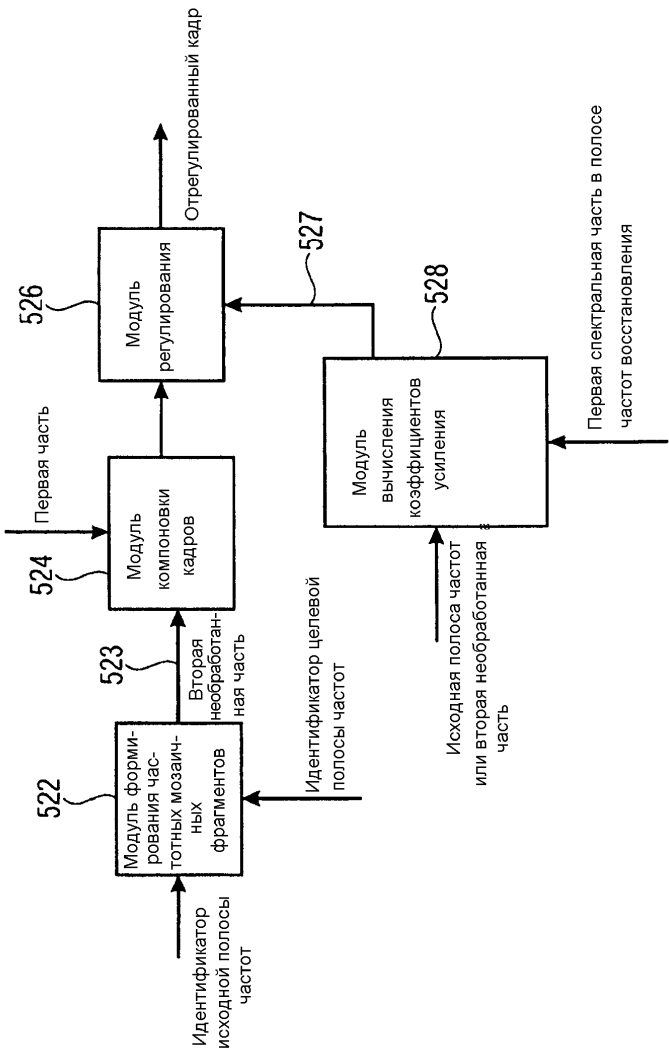


(Другие спектральные части)

Фиг. 5a

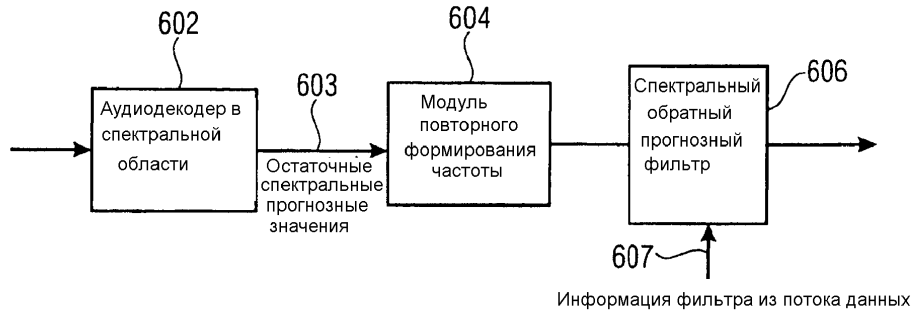


Фиг. 5b

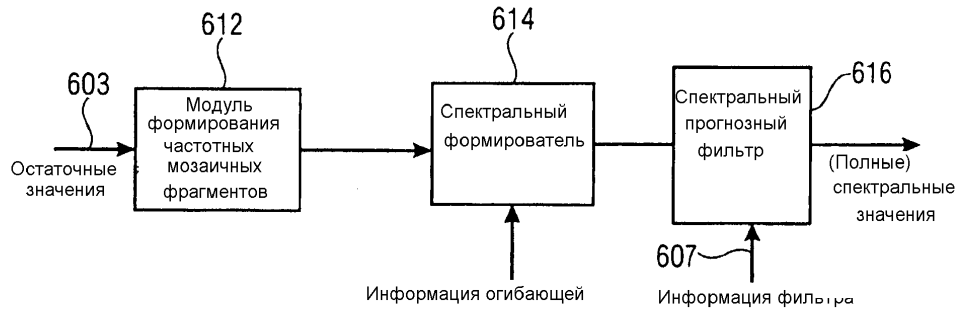


Фиг. 5с

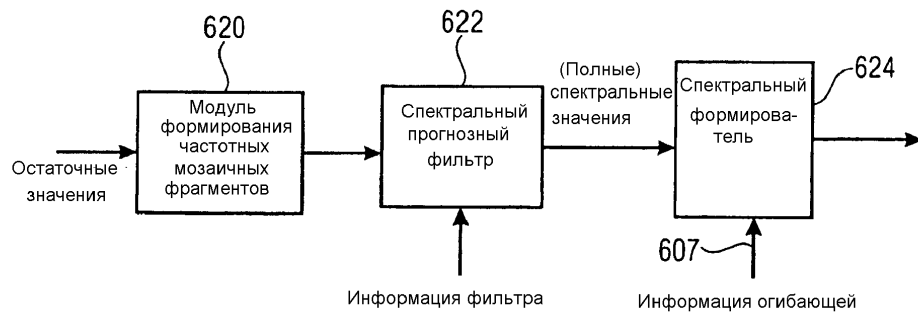
10/29



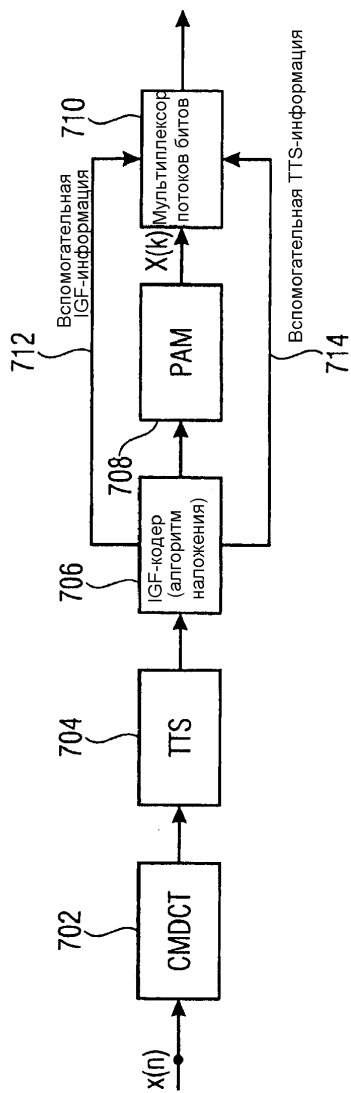
Фиг.6а



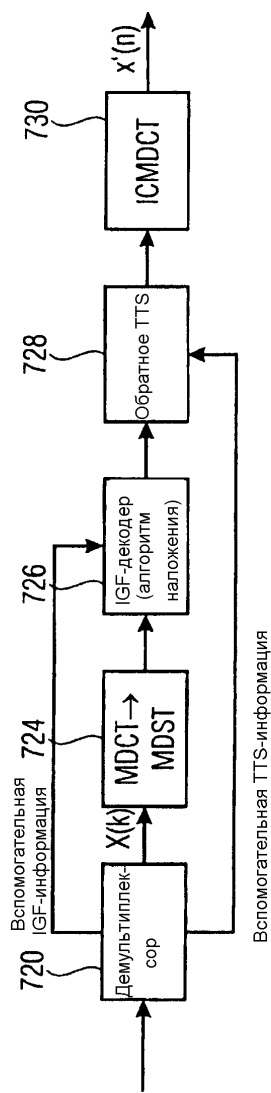
Фиг.6b



Фиг.6с

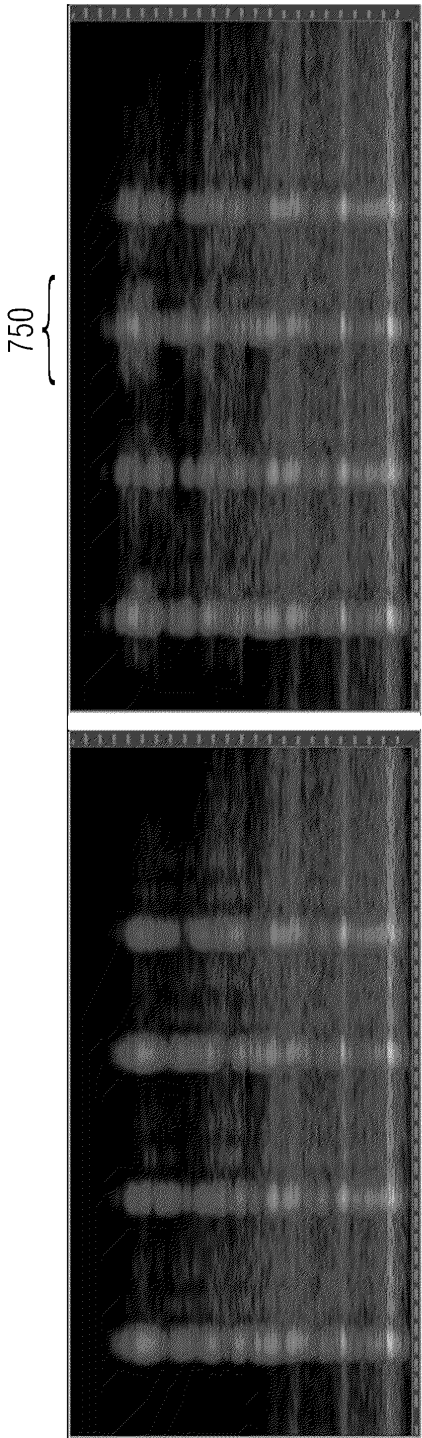


(Кодер)
Фиг. 7а



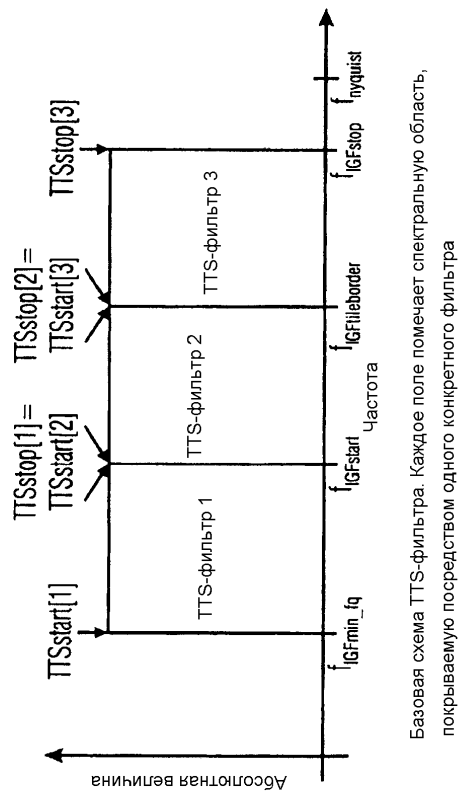
(Декодер)
Фиг. 7б

12/29



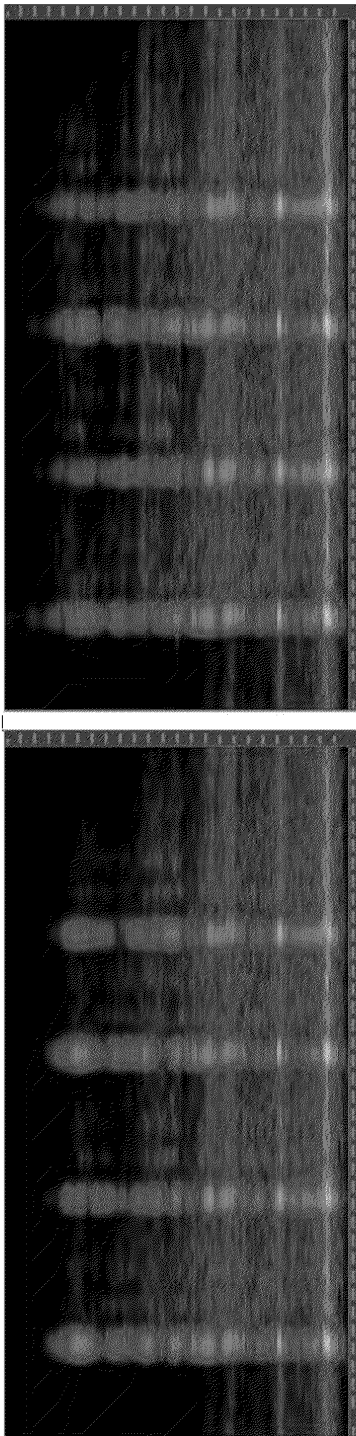
DFT-спектрограмма исходного сигнала (левая панель);
DFT-спектрограмма расширенного сигнала без TTS (правая панель)

Фиг. 7с



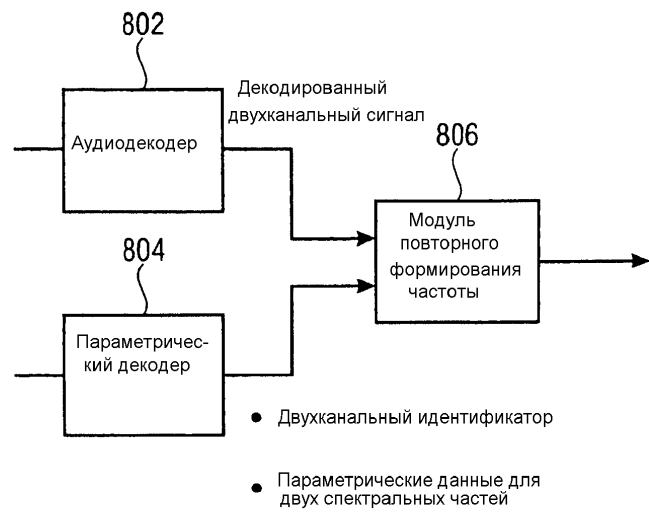
Фиг. 7d

750



DFT-спектрограмма исходного сигнала (левая панель);
DFT-спектрограмма расширенного сигнала с TTS (правая панель)

Фиг. 7е

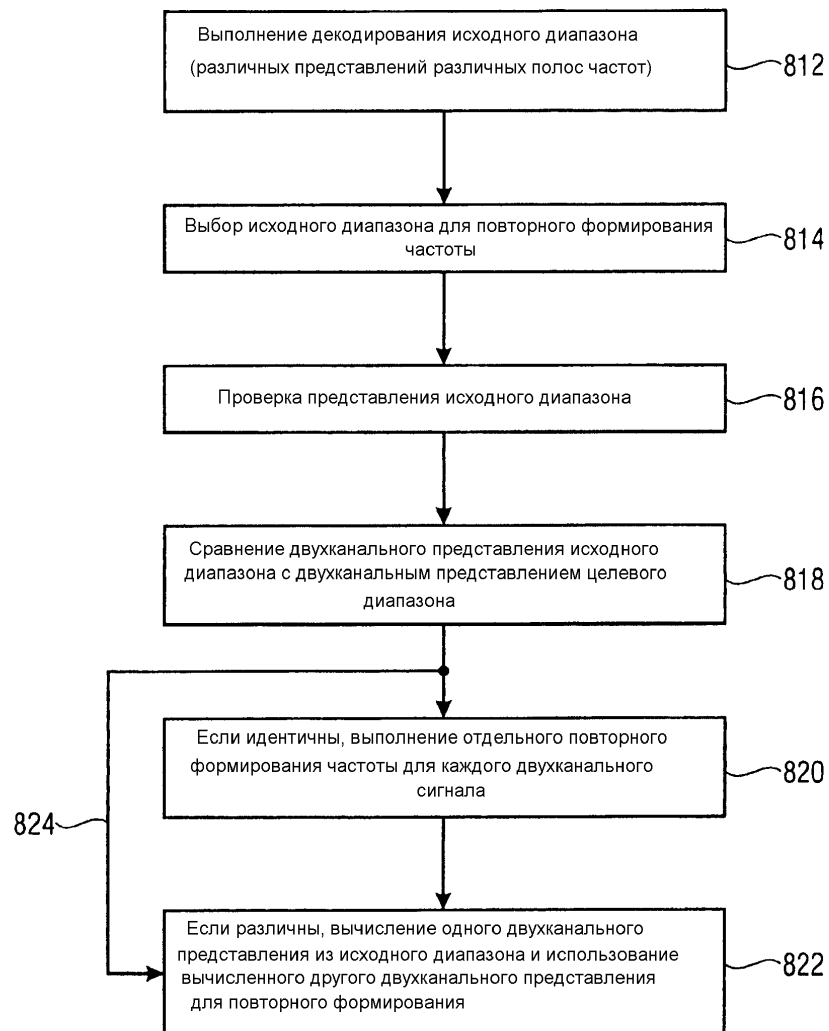


Фиг. 8a

| | | | | | |
|-------------------|---|---|---|---|--|
| Исходный диапазон | 1 | 1 | 2 | 2 | 1: Первое двухканальное представление (к примеру, L/R) |
| Целевой диапазон | 1 | 2 | 1 | 2 | 2: Второе двухканальное представление (к примеру, M/S) |

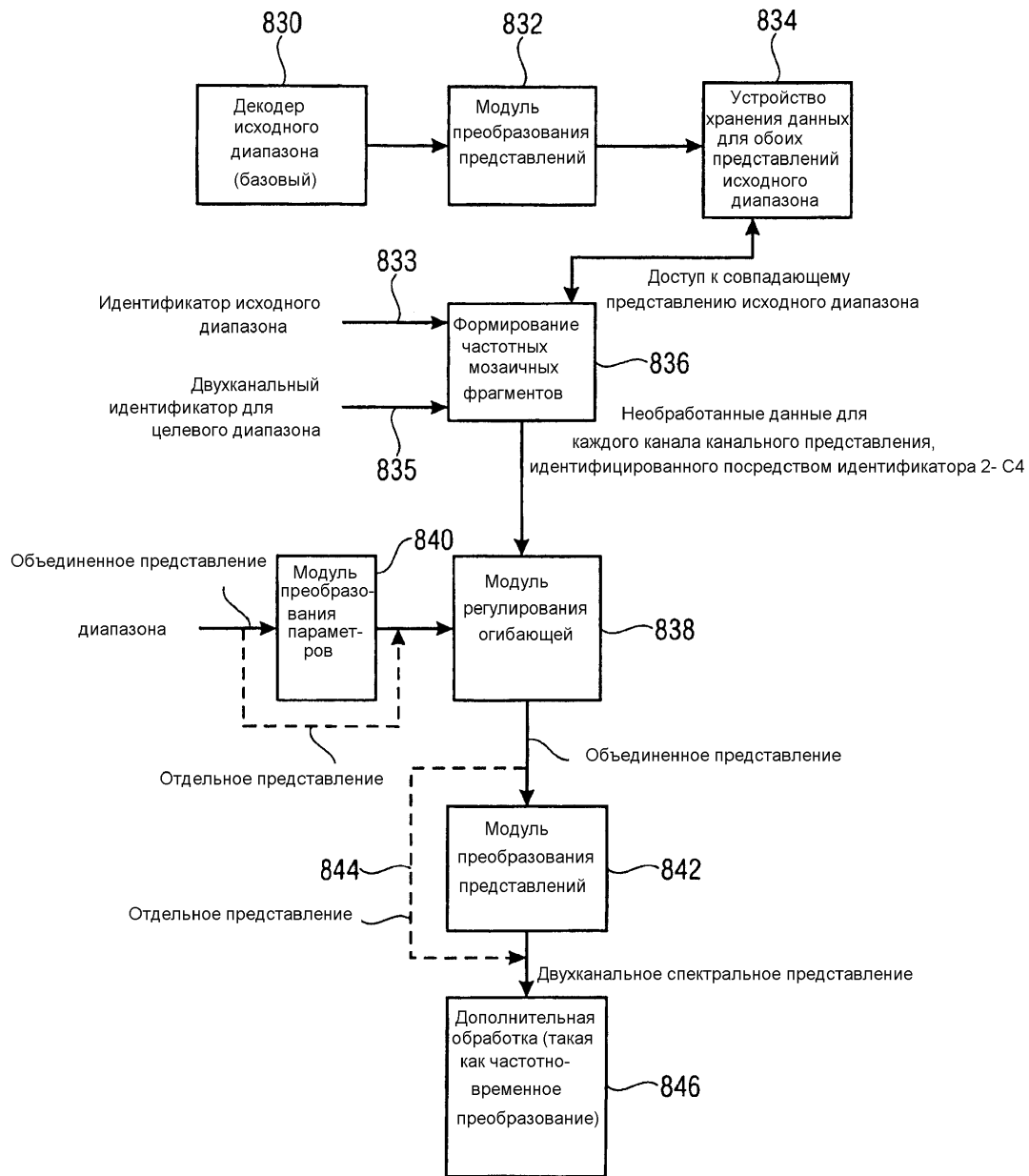
Фиг. 8b

16/29

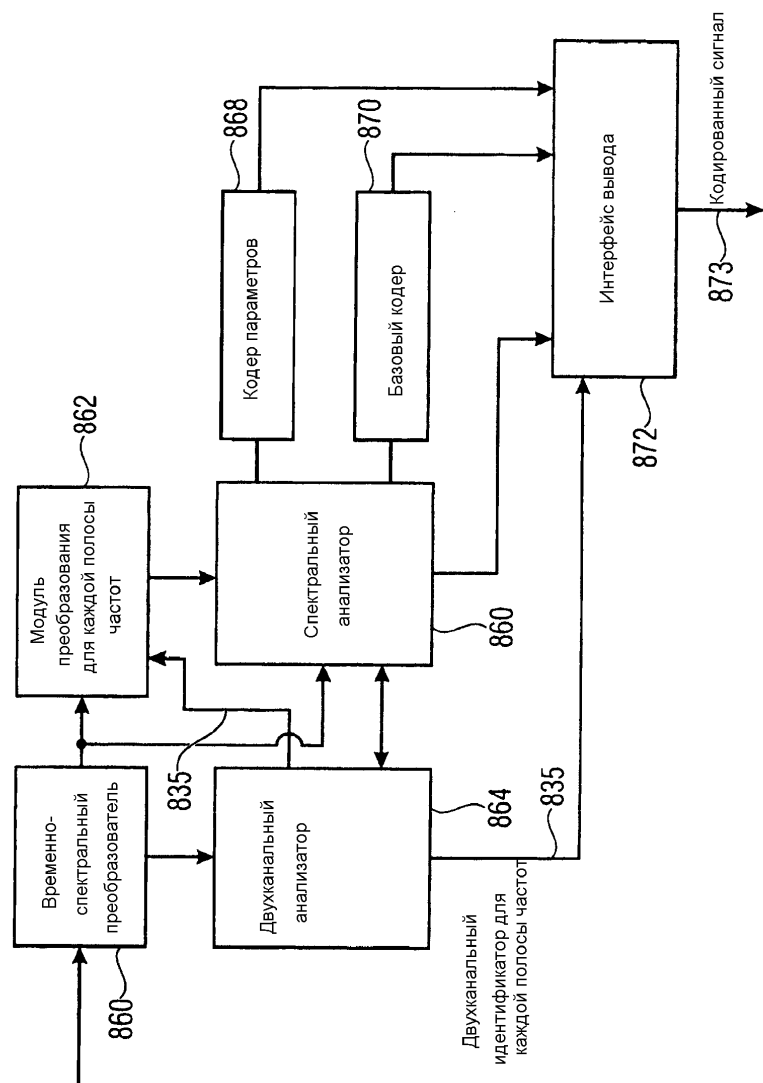


Фиг. 8с

17/29

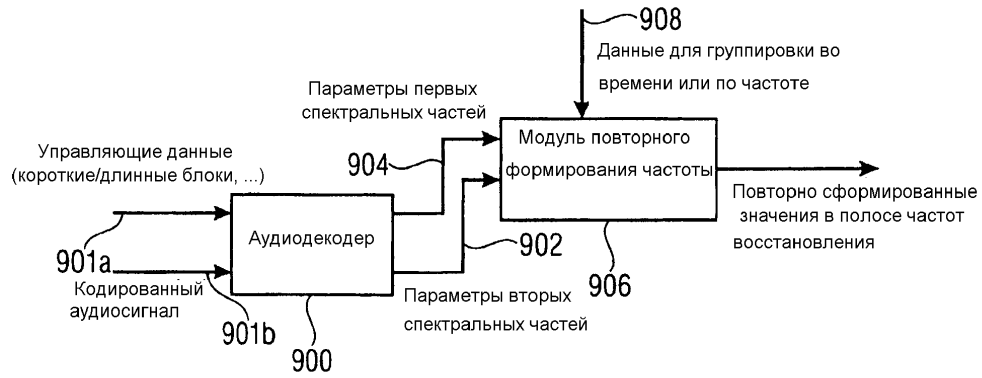


Фиг.8d

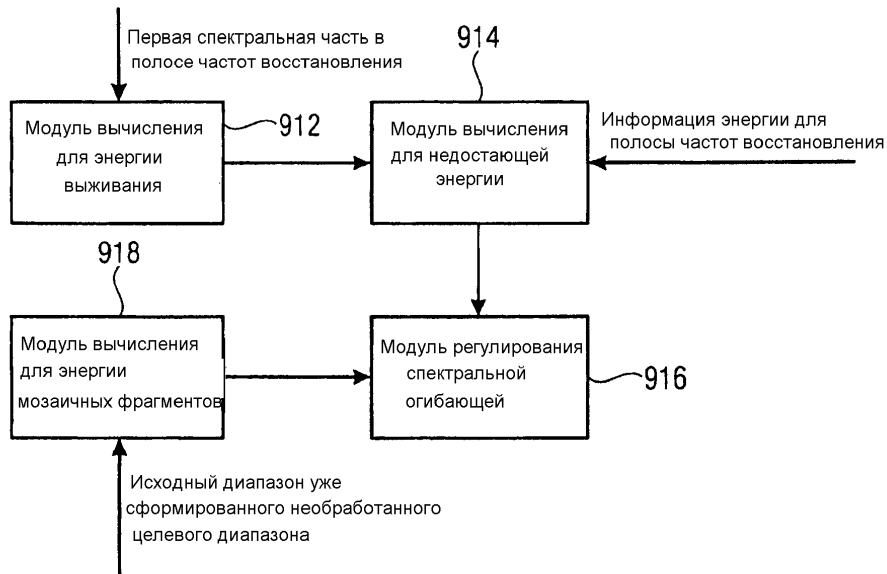


Фиг. 8е

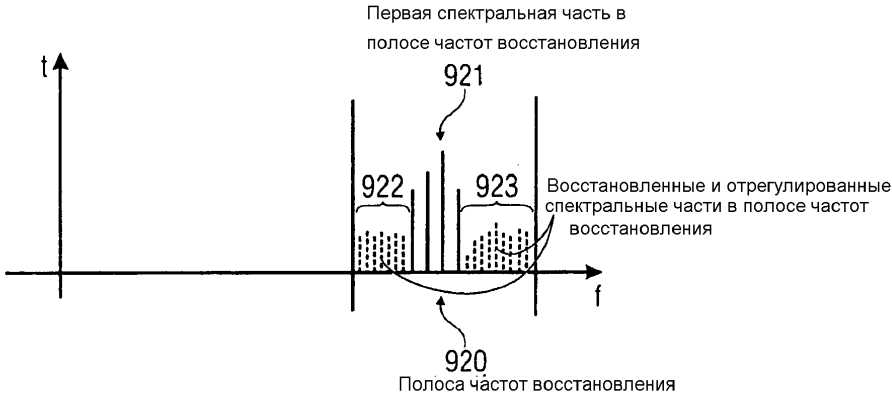
19/29



Фиг.9а



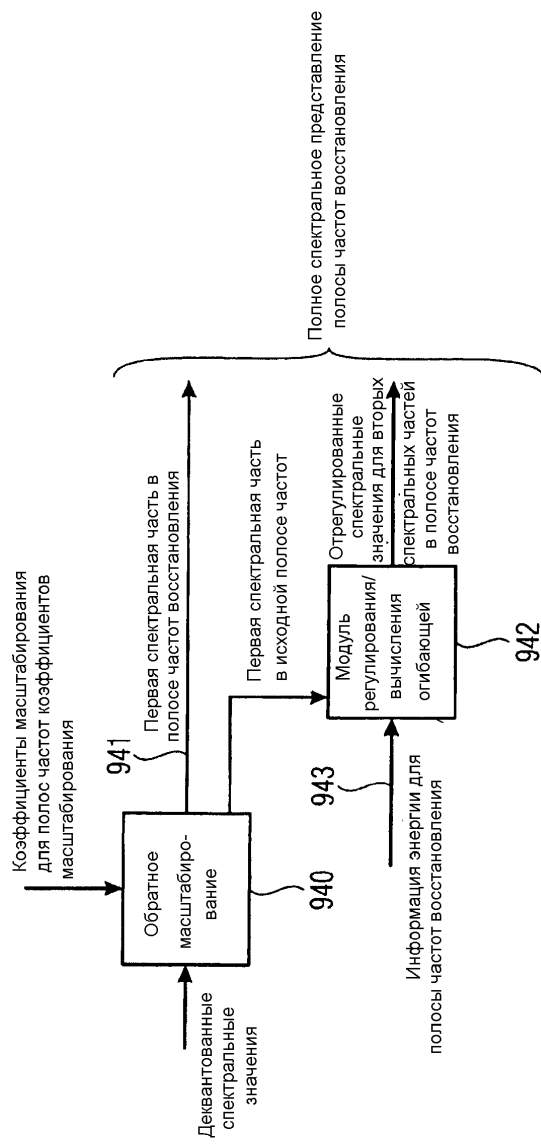
Фиг.9б



- К примеру:
- Энергия выживания 5 единиц
 - Значение энергии для полосы частот восстановления 10 единиц (покрывает первую и вторую спектральные части в полосе частот восстановления)
 - Энергия данных исходного диапазона или необработанных данных целевого диапазона 8 единиц
 - Недостаточная энергия 5 единиц
 - Коэффициент усиления $g := \sqrt{\frac{mE_k}{pE_k}} = 0.79$

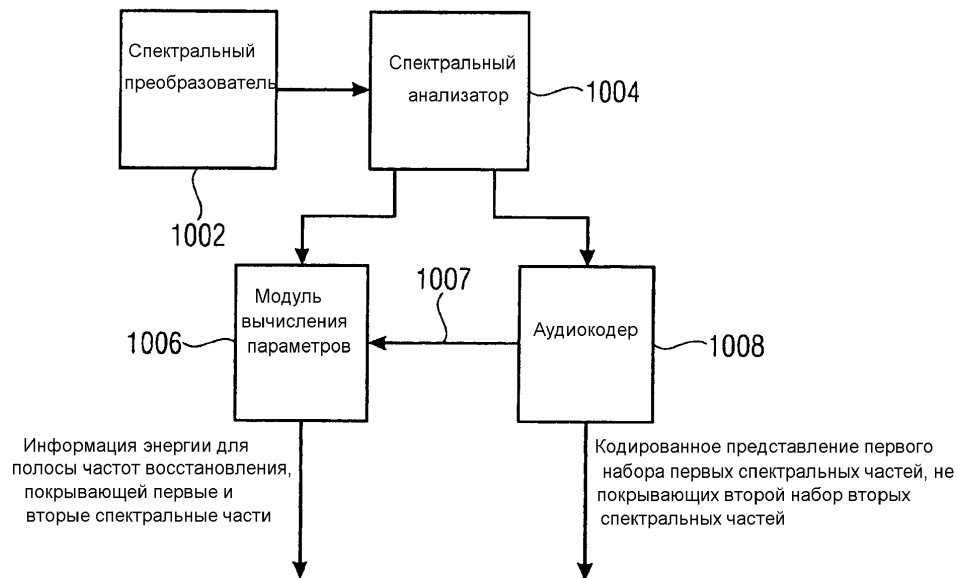
- Регулируются только спектральные значения для вторых спектральных частей
- Первая спектральная часть не затрагивается посредством регулирования огибающей

Фиг.9с



Фиг. 9d

22/29

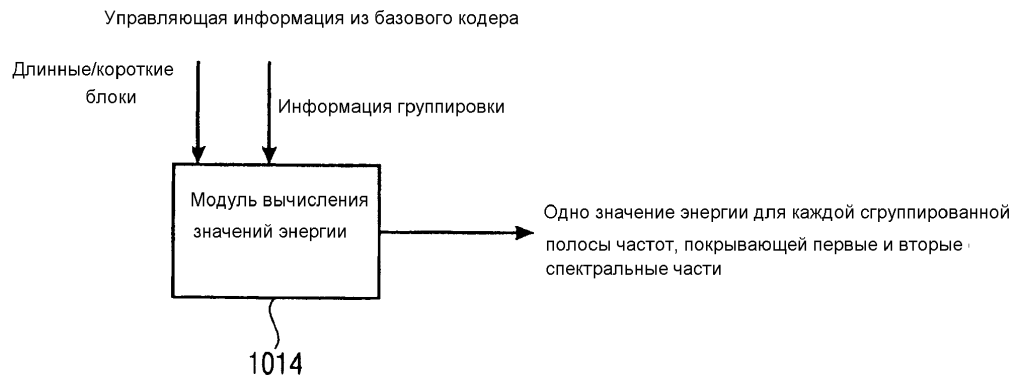


Фиг. 10a

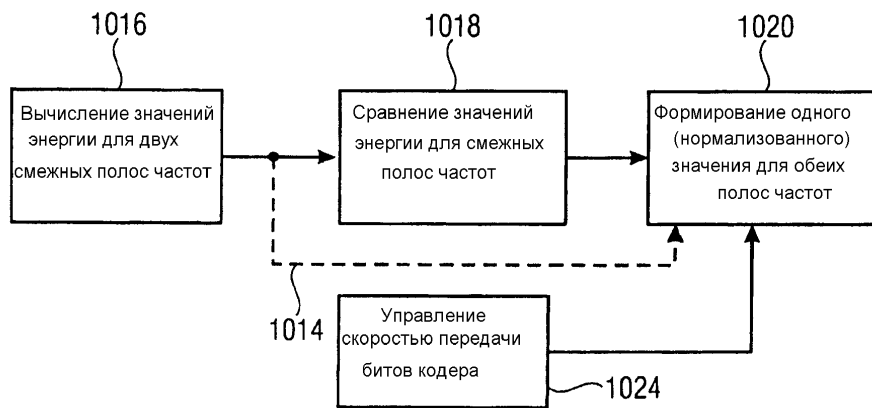


Фиг. 10b

23/29

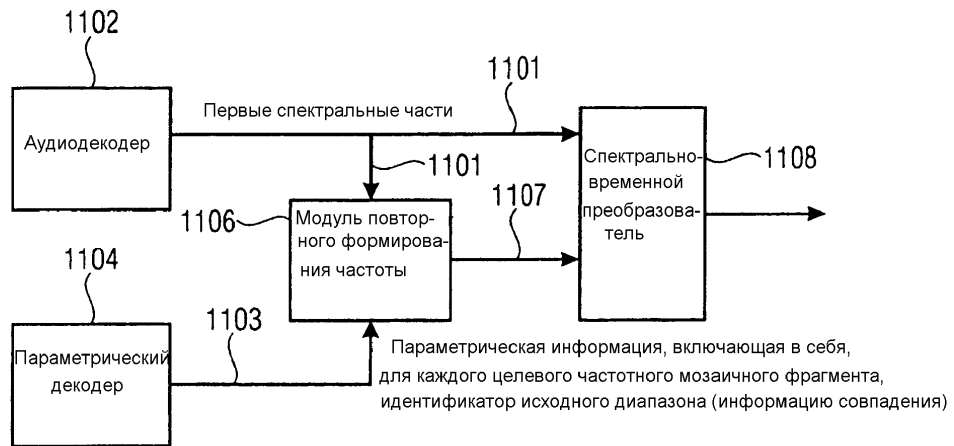


Фиг.10с

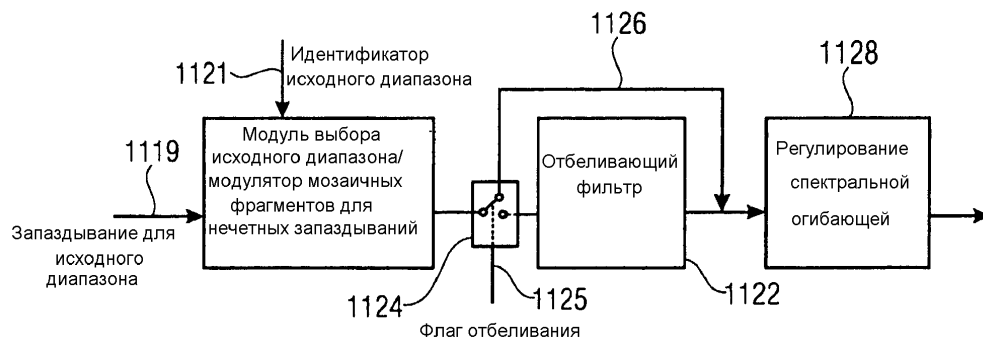


Фиг.10d

24/29

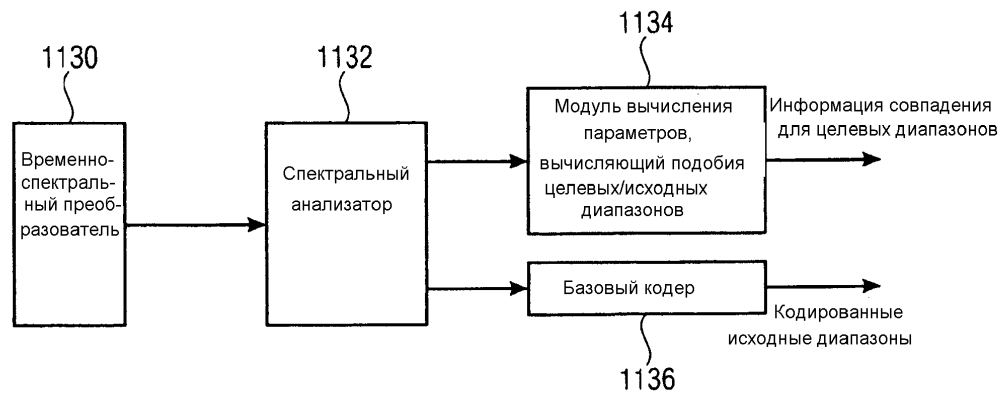


Фиг.11а

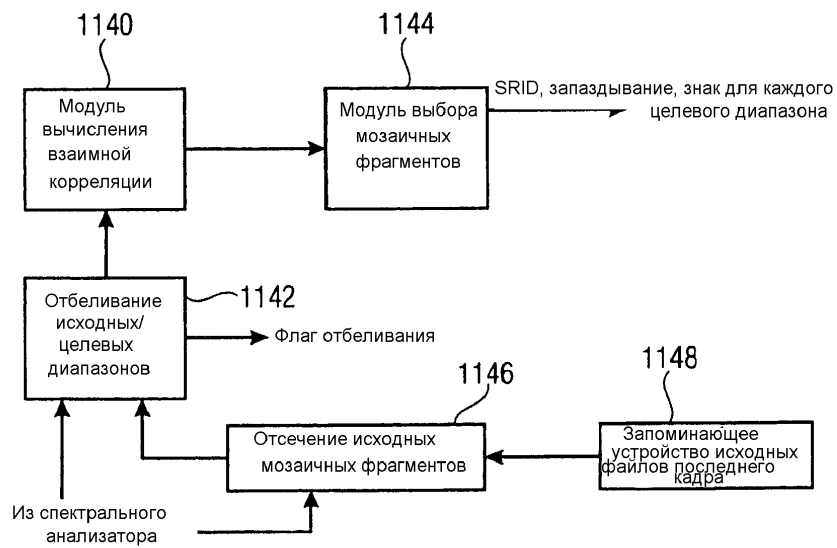


Фиг.11б

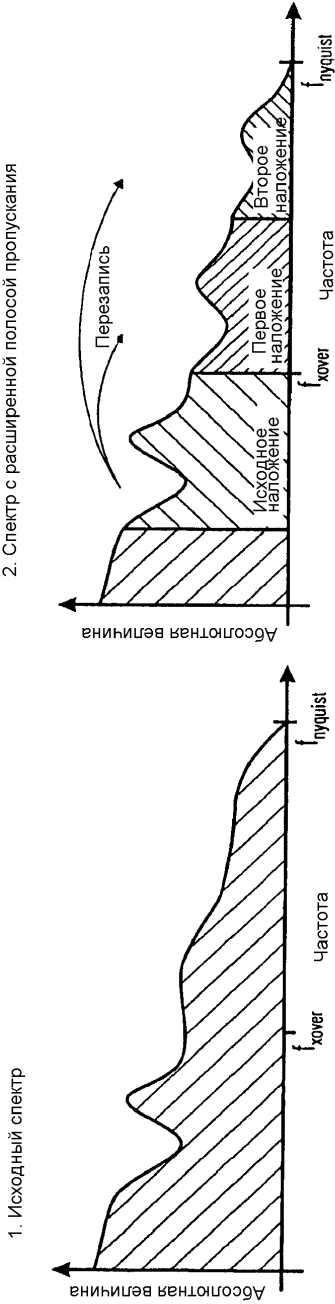
25/29



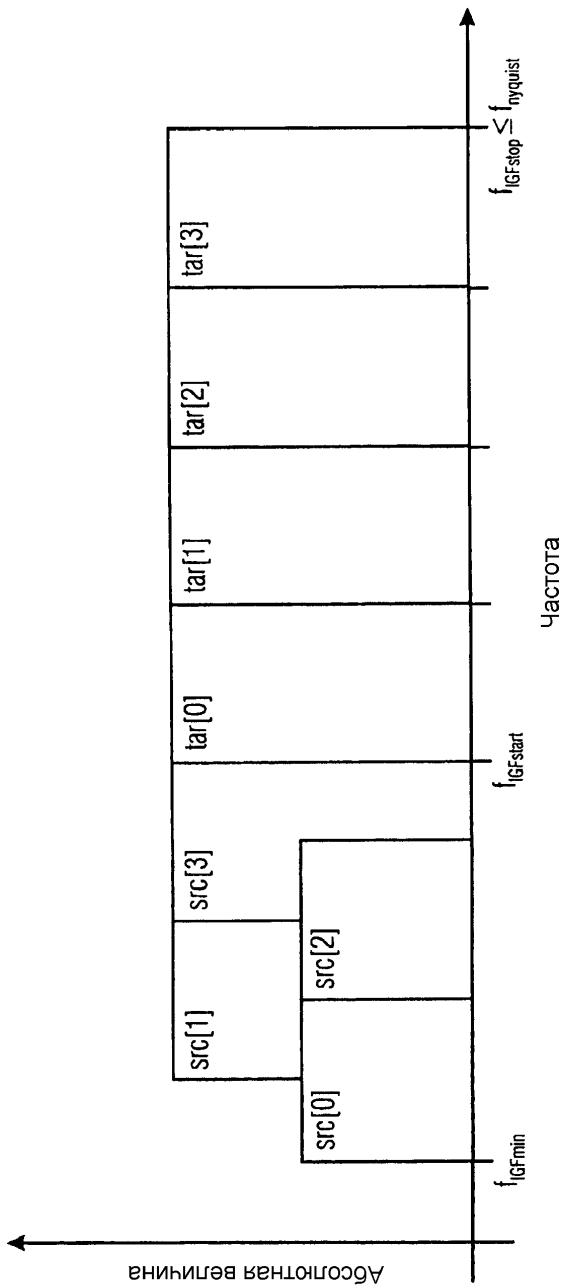
Фиг.11с



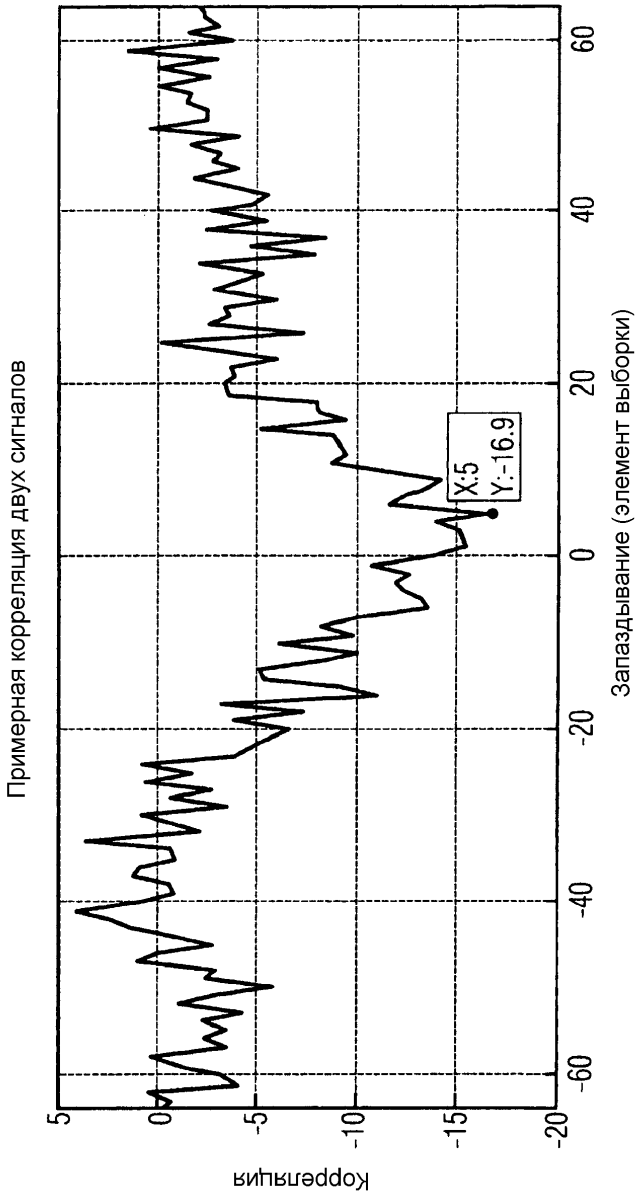
Фиг.11d



Фиг. 12а

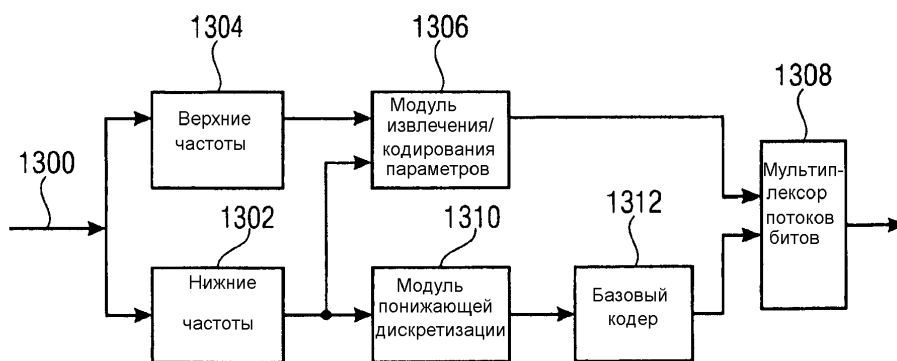


Фиг. 12b



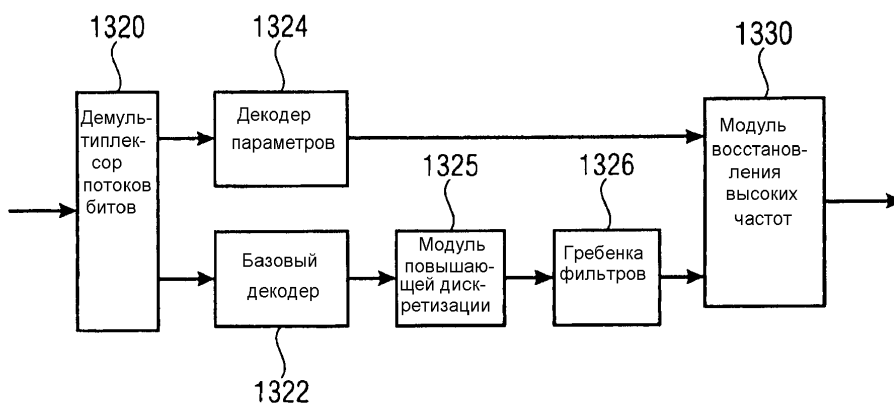
Фиг. 12с

29/29



(Предшествующий уровень техники)

Фиг. 13a



(Предшествующий уровень техники)

Фиг. 13b