



(51) МПК
G10L 19/005 (2013.01)
G10L 19/04 (2013.01)
G10L 19/12 (2013.01)
G10L 19/22 (2013.01)

**ФЕДЕРАЛЬНАЯ СЛУЖБА
 ПО ИНТЕЛЛЕКТУАЛЬНОЙ СОБСТВЕННОСТИ**

(12) ОПИСАНИЕ ИЗОБРЕТЕНИЯ К ПАТЕНТУ

(21)(22) Заявка: 2015122128, 13.11.2013

(24) Дата начала отсчета срока действия патента:
 13.11.2013

Дата регистрации:
 13.09.2017

Приоритет(ы):

(30) Конвенционный приоритет:
 13.11.2012 US 61/725,694

(43) Дата публикации заявки: 10.01.2017 Бюл. № 1

(45) Опубликовано: 13.09.2017 Бюл. № 26

(85) Дата начала рассмотрения заявки РСТ на
 национальной фазе: 15.06.2015

(86) Заявка РСТ:
 KR 2013/010310 (13.11.2013)

(87) Публикация заявки РСТ:
 WO 2014/077591 (22.05.2014)

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

(72) Автор(ы):

**ЧОО Ки-хиун (KR),
 Поров Антон Викторович (RU),
 Осипов Константин Сергеевич (RU),
 ЛИ Нам-сук (KR)**

(73) Патентообладатель(и):

**САМСУНГ ЭЛЕКТРОНИКС КО., ЛТД.
 (KR)**

(56) Список документов, цитированных в отчете
 о поиске: US 2012/0069899 A1, 22.03.2012. EP
 2096629 A1, 02.09.2009. US 2012/0253797 A,
 04.10.2012. US 2005/0256701 A1, 17.11.2005. US
 2007/0179783 A1, 02.08.2007. EP 2144230 A1,
 13.01.2010. RU 2428748 C2, 10.09.2011.

**(54) СПОСОБ И УСТРОЙСТВО ДЛЯ ОПРЕДЕЛЕНИЯ РЕЖИМА КОДИРОВАНИЯ, СПОСОБ И
 УСТРОЙСТВО ДЛЯ КОДИРОВАНИЯ АУДИОСИГНАЛОВ И СПОСОБ И УСТРОЙСТВО ДЛЯ
 ДЕКОДИРОВАНИЯ АУДИОСИГНАЛОВ**

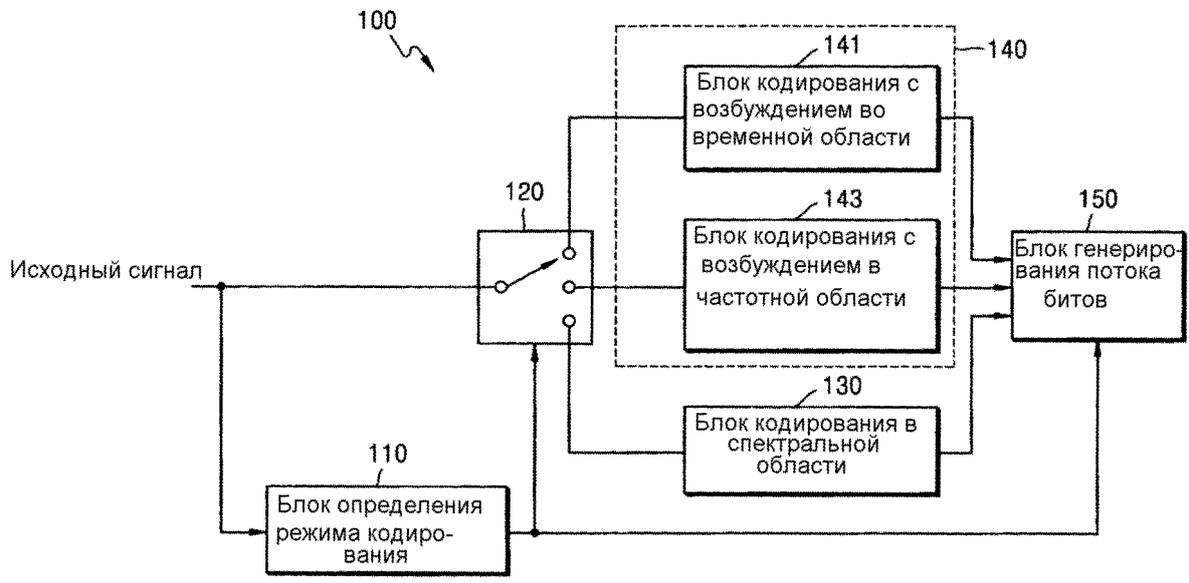
(57) Реферат:

Изобретение относится к средствам для определения режима кодирования и для кодирования/декодирования аудиосигналов. Технический результат заключается в сокращении задержек, вызванных частым изменением режима кодирования. Способ определения режима кодирования включает в себя определение одного из множества режимов кодирования, включающего в себя первый режим кодирования

и второй режим кодирования, в качестве первоначального режима кодирования в соответствии с характеристиками аудиосигнала и, если имеется ошибка в определении первоначального режима кодирования, генерирование скорректированного режима кодирования путем коррекции первоначального режима кодирования на третий режим кодирования. 1 з.п. ф-лы, 9 ил.

RU 2 630 889 C2

RU 2 630 889 C2



ФИГ.1

RU 2630889 C2

RU 2630889 C2



FEDERAL SERVICE
FOR INTELLECTUAL PROPERTY

(51) Int. Cl.
G10L 19/005 (2013.01)
G10L 19/04 (2013.01)
G10L 19/12 (2013.01)
G10L 19/22 (2013.01)

(12) **ABSTRACT OF INVENTION**

(21)(22) Application: **2015122128**, 13.11.2013

(24) Effective date for property rights:
13.11.2013

Registration date:
13.09.2017

Priority:

(30) Convention priority:
13.11.2012 US 61/725,694

(43) Application published: **10.01.2017** Bull. № 1

(45) Date of publication: **13.09.2017** Bull. № 26

(85) Commencement of national phase: **15.06.2015**

(86) PCT application:
KR 2013/010310 (13.11.2013)

(87) PCT publication:
WO 2014/077591 (22.05.2014)

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

(72) Inventor(s):

**CHOO Ki-khiun (KR),
Porov Anton Viktorovich (RU),
Osipov Konstantin Sergeevich (RU),
LI Nam-suk (KR)**

(73) Proprietor(s):

SAMSUNG ELEKTRONIKS KO., LTD. (KR)

(54) **METHOD AND DEVICE FOR DETERMINING THE CODING MODE, METHOD AND DEVICE FOR CODING AUDIO SIGNALS AND A METHOD AND DEVICE FOR DECODING AUDIO SIGNALS**

(57) Abstract:

FIELD: physics.

SUBSTANCE: encoding mode determination method includes determining one of a plurality of coding modes including the first coding mode and the second coding mode as the initial coding mode in accordance with the audio signal characteristics and if there is an error in determining the initial coding mode,

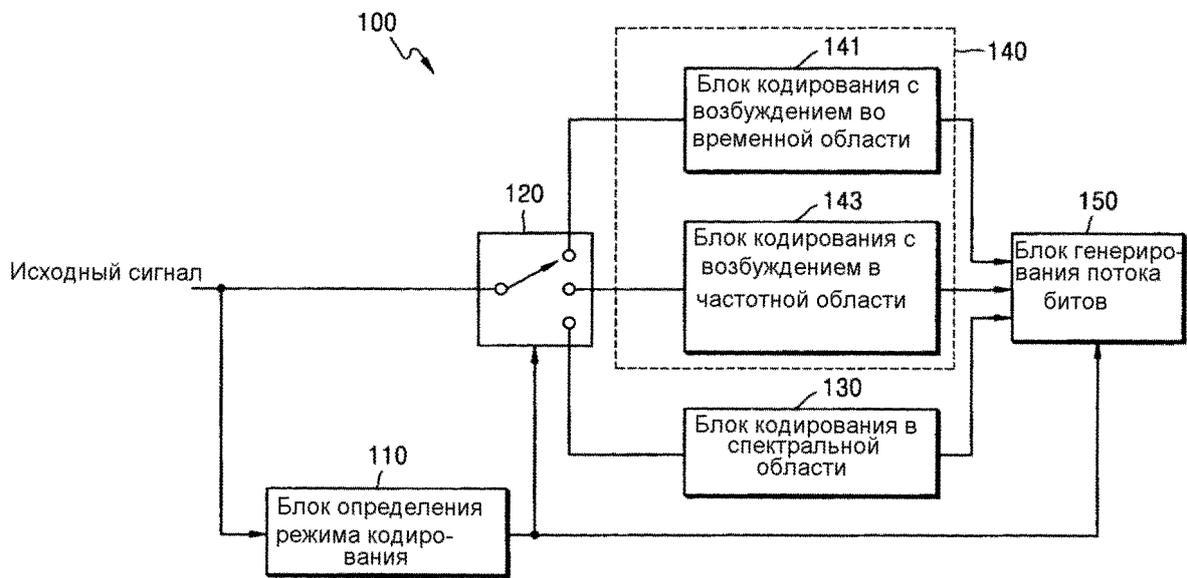
generating the corrected the encoding mode by correcting the initial encoding mode to the third encoding mode.

EFFECT: reducing delays caused by frequent change of coding mode.

2 cl, 9 dwg

C 2
2 6 3 0 8 8 9
R U

R U
2 6 3 0 8 8 9
C 2



ФИГ.1

ОБЛАСТЬ ТЕХНИКИ

[1] Устройства и способы, согласующиеся с примерными вариантами осуществления, относятся к аудио - кодированию и декодированию, и более конкретно к способу и устройству для определения режима кодирования для улучшения качества реконструированного аудиосигнала путем определения режима кодирования, подходящего к характеристикам аудиосигнала, и предотвращения частого переключения режимов кодирования, к способу и устройству для кодирования аудиосигнала, а также к способу и устройству для декодирования аудиосигнала.

УРОВЕНЬ ТЕХНИКИ

[2] Широко известно, что эффективно кодировать музыкальный сигнал в частотной области, и эффективно кодировать речевой сигнал во временной области. Поэтому были предложены различные методы для определения класса аудиосигнала, в котором смешаны музыкальный сигнал и речевой сигнал, а также для определения режима кодирования в соответствии с определенным классом.

[3] Однако из-за частоты переключения режимов кодирования происходят не только задержки, но также ухудшается и качество декодированного звука. Кроме того, поскольку не существует никакого метода для корректировки первоначально определенного режима кодирования, то есть класса, если во время определения режима кодирования случается ошибка, качество реконструированного аудиосигнала ухудшается.

СУЩНОСТЬ ИЗОБРЕТЕНИЯ ТЕХНИЧЕСКАЯ ЗАДАЧА

[4] Аспекты одного или более примерных вариантов осуществления обеспечивают способ и устройство для определения режима кодирования для улучшения качества реконструированного аудиосигнала путем определения режима кодирования, подходящего к характеристикам аудиосигнала, способ и устройство для кодирования аудиосигнала, а также способ и устройство для декодирования аудиосигнала.

[5] Аспекты одного или более примерных вариантов осуществления обеспечивают способ и устройство для определения режима кодирования, подходящего к характеристикам аудиосигнала, и сокращения задержек вследствие частого переключения режимов кодирования, способ и устройство для кодирования аудиосигнала, а также способ и устройство для декодирования аудиосигнала.

ТЕХНИЧЕСКОЕ РЕШЕНИЕ

[6] В соответствии с одним аспектом одного или более примерных вариантов осуществления имеется способ определения режима кодирования, включающий в себя определение одного из множества режимов кодирования, включающего в себя первый режим кодирования и второй режим кодирования, в качестве первоначального режима кодирования в соответствии с характеристиками аудиосигнала, и, если имеется ошибка в определении первоначального режима кодирования, генерирование скорректированного режима кодирования путем коррекции первоначального режима кодирования на третий режим кодирования.

[7] В соответствии с одним аспектом одного или более примерных вариантов осуществления предлагается способ кодирования аудиосигнала, включающий в себя определение одного из множества режимов кодирования, включающего в себя первый режим кодирования и второй режим кодирования в качестве первоначального режима кодирования в соответствии с характеристиками аудиосигнала, и, если имеется ошибка в определении первоначального режима кодирования, генерирование скорректированного режима кодирования путем коррекции первоначального режима

кодирования на третий режим кодирования, а также выполнение различных процессов кодирования над аудиосигналом на основе либо первоначального режима кодирования, либо скорректированного режима кодирования.

5 [8] В соответствии с одним аспектом одного или более примерных вариантов осуществления предлагается способ декодирования аудиосигнала, включающий в себя анализ потока битов, содержащего одно из первоначального режима кодирования, полученного путем определения одного из множества режимов кодирования, включающего в себя первый режим кодирования и второй режим кодирования в соответствии с характеристиками аудиосигнала, а также третий режим кодирования, 10 скорректированный из первоначального режима кодирования, если имеется ошибка в определении первоначального режима кодирования, а также выполнение различных процессов декодирования над потоком битов на основе либо первоначального режима кодирования, либо третьего режима кодирования.

ПОЛЕЗНЫЕ ЭФФЕКТЫ

15 [9] В соответствии с примерными вариантами осуществления, путем определения окончательного режима кодирования текущего кадра на основе коррекции первоначального режима кодирования и режимов кодирования кадров, соответствующих продолжительности затягивания ("hangover"), может быть выбран режим кодирования, адаптирующийся к характеристикам аудиосигнала, предотвращая 20 частое переключение режимов кодирования между кадрами.

КРАТКОЕ ОПИСАНИЕ ЧЕРТЕЖЕЙ

[10] Фиг. 1 представляет собой структурную схему, иллюстрирующую конфигурацию устройства аудиокодирования в соответствии с одним примерным вариантом осуществления;

25 [11] Фиг. 2 представляет собой структурную схему, иллюстрирующую конфигурацию устройства аудиокодирования в соответствии с другим примерным вариантом осуществления;

[12] Фиг. 3 представляет собой структурную схему, иллюстрирующую конфигурацию блока определения режима кодирования в соответствии с одним примерным вариантом 30 осуществления;

[13] Фиг. 4 представляет собой структурную схему, иллюстрирующую конфигурацию блока определения первоначального режима кодирования в соответствии с одним примерным вариантом осуществления;

35 [14] Фиг. 5 представляет собой структурную схему, иллюстрирующую конфигурацию блока извлечения параметров особенности в соответствии с одним примерным вариантом осуществления;

[15] Фиг. 6 представляет собой схему, иллюстрирующую способ адаптивного переключения между кодированием в области линейного предсказания и спектральной областью в соответствии с одним примерным вариантом осуществления;

40 [16] Фиг. 7 представляет собой схему, иллюстрирующую работу блока корректировки режима кодирования в соответствии с одним примерным вариантом осуществления;

[17] Фиг. 8 представляет собой структурную схему, иллюстрирующую конфигурацию устройства аудиодекодирования в соответствии с одним примерным вариантом осуществления; и

45 [18] Фиг. 9 представляет собой структурную схему, иллюстрирующую конфигурацию устройства аудиодекодирования в соответствии с другим примерным вариантом осуществления.

ВАРИАНТЫ ОСЩЕСТВЛЕНИЯ ИЗОБРЕТЕНИЯ

[19] Далее будут подробно рассмотрены варианты осуществления, примеры которых проиллюстрированы на сопроводительных чертежах, при этом на всех из них одинаковые ссылочные позиции относятся к одинаковым элементам. В этой связи представленные варианты осуществления могут иметь различные формы и не должны рассматриваться как ограниченные описаниями, сформулированными в настоящем документе. Соответственно, эти варианты осуществления просто описываются ниже со ссылками на фигуры для того, чтобы пояснить аспекты настоящего описания.

[20] Такие термины, как «соединенный» и «связанный» могут быть использованы для указания состояния непосредственного соединения и связи, но следует понимать, что между ними может быть размещен другой компонент.

[21] Такие термины, как «первый» и «второй» могут быть использованы для описания различных компонентов, но эти компоненты не должны ограничиваться этими терминами. Эти термины могут быть использованы, чтобы отличать один компонент от другого компонента.

[22] Блоки, описанные в примерных вариантах осуществления, проиллюстрированы независимо для указания различных характерных функции, и это не означает, что каждый блок образован из одного отдельного аппаратного или программного компонента. Каждый блок проиллюстрирован для удобства объяснения, и множество блоков могут образовывать один блок, и один блок может быть разделен на множество блоков.

[23] Фиг. 1 представляет собой структурную схему, иллюстрирующую конфигурацию устройства 100 аудиокодирования в соответствии с одним примерным вариантом осуществления.

[24] Устройство 100 аудиокодирования, показанное на Фиг. 1, может включать в себя блок 110 определения режима кодирования, переключающий блок 120, блок 130 кодирования в спектральной области, блок 140 кодирования в области линейного предсказания, а также блок 150 генерирования потока битов. Блок 140 кодирования в области линейного предсказания может включать в себя блок 141 кодирования с возбуждением во временной области и блок 143 кодирования с возбуждением в частотной области, где блок 140 кодирования в области линейного предсказания может быть воплощен как по меньшей мере один из двух блоков 141 и 143 кодирования с возбуждением. Если нет необходимости воплощения в виде отдельного аппаратного обеспечения, вышеперечисленные компоненты могут быть интегрированы в по меньшей мере один модуль и могут быть осуществлены как по меньшей мере один процессор (не показан). Здесь, термин аудиосигнал может относиться к музыкальному сигналу, речевому сигналу или к их смешанному сигналу.

[25] Со ссылкой на Фиг. 1, блок 110 определения режима кодирования может анализировать характеристики аудиосигнала для того, чтобы определить класс аудиосигнала и определить режим кодирования в соответствии с результатом этой классификации. Определение режима кодирования может быть выполнено в единицах суперкадров, кадров или полос. Альтернативно, определение режима кодирования может быть выполнено в единицах множества групп суперкадров, множества групп кадров или множества групп полос. В настоящем документе примеры режимов кодирования могут включать в себя спектральную область и временную область или область линейного предсказания, но не ограничиваются этим. Если характеристики и скорость обработки процессора являются достаточными, и задержки из-за переключения режимов кодирования могут быть устранены, режимы кодирования могут быть подразделены, и схемы кодирования также могут быть подразделены в соответствии

с режимом кодирования. В соответствии с одним примерным вариантом осуществления блок 110 определения режима кодирования может определять первоначальный режим кодирования аудиосигнала как один из режима кодирования в спектральной области и режима кодирования во временной области. В соответствии с другим примерным

5 вариантом осуществления блок 110 определения режима кодирования может определять первоначальный режим кодирования аудиосигнала как один из режима кодирования в спектральной области, режима кодирования с возбуждением во временной области и режима кодирования с возбуждением в частотной области. Если режим кодирования в спектральной области определяется как первоначальный режим кодирования, блок

10 110 определения режима кодирования может скорректировать первоначальный режим кодирования на один из режима кодирования в спектральной области и режима кодирования с возбуждением в частотной области. Если режим кодирования во временной области, то есть режим кодирования с возбуждением во временной области, определяется как первоначальный режим кодирования, блок 110 определения режима кодирования может скорректировать первоначальный режим кодирования на один из

15 режима кодирования с возбуждением во временной области и режима кодирования с возбуждением в частотной области. Если режим кодирования с возбуждением во временной области определяется как первоначальный режим кодирования, определение окончательного режима кодирования может быть выполнено по выбору. Другими

20 словами, первоначальный режим кодирования, то есть режим кодирования с возбуждением во временной области, может быть сохранен. Блок 110 определения режима кодирования может определять режимы кодирования множества кадров, соответствующих продолжительности затягивания, а также может определять окончательный режим кодирования для текущего кадра. В соответствии с одним

25 примерным вариантом осуществления, если первоначальный режим кодирования или скорректированный режим кодирования текущего кадра идентичен режимам кодирования множества предыдущих кадров, например 7 предыдущих кадров, соответствующий первоначальный режим кодирования или скорректированный режим кодирования могут быть определены как окончательный режим кодирования текущего кадра. Тем временем, если первоначальный режим кодирования или скорректированный режим кодирования текущего кадра не идентичен режимам кодирования множества

30 предыдущих кадров, например, 7 предыдущих кадров, блок 110 определения режима кодирования может определить режим кодирования кадра, непосредственно перед текущим кадром, как окончательный режим кодирования текущего кадра.

35 [26] Как было описано выше, путем определения окончательного режима кодирования текущего кадра на основе коррекции первоначального режима кодирования и режимов кодирования кадров, соответствующих продолжительности затягивания, может быть выбран режим кодирования, адаптирующийся к характеристикам аудиосигнала, предотвращая частое переключение режимов

40 кодирования между кадрами.

[27] Обычно кодирование во временной области, то есть, кодирование с возбуждением во временной области, может быть эффективным для речевого сигнала, кодирование в спектральной области может быть эффективным для музыкального сигнала, и кодирование с возбуждением в частотной области может быть эффективным для

45 вокального и/или гармонического сигнала.

[28] В соответствии с режимом кодирования, определенным блоком 110 определения режима кодирования, переключающий блок 120 может обеспечить аудиосигнал либо к блоку 130 кодирования в спектральной области, либо к блоку 140 кодирования в

области линейного предсказания. Если блок 140 кодирования в области линейного предсказания воплощается как блок 141 кодирования с возбуждением во временной области, переключающий блок 120 может включать в себя всего две ветви. Если блок 140 кодирования в области линейного предсказания воплощается как блок 141 кодирования с возбуждением во временной области и блок 143 кодирования с возбуждением в частотной области, переключающий блок 120 может иметь всего 3 ветви.

[29] Блок 130 кодирования в спектральной области может кодировать аудиосигнал в спектральной области. Спектральная область может относиться к частотной области или к области преобразования. Примеры способов кодирования, применимых к блоку 130 кодирования в спектральной области, могут включать в себя усовершенствованное аудиокодирование (AAC) или комбинацию модифицированного дискретного косинусного преобразования (MDCT) и факториального импульсного кодирования (FPC), но не ограничивается этим. Более подробно, другие методы квантования и методы энтропийного кодирования могут использоваться вместо FPC. Может быть эффективным кодировать музыкальный сигнал в блоке 130 кодирования в спектральной области.

[30] Блок 140 кодирования в области линейного предсказания может кодировать аудиосигнал в области линейного предсказания. Область линейного предсказания может относиться к области возбуждения или к временной области. Блок 140 кодирования в области линейного предсказания может быть воплощен как блок 141 кодирования с возбуждением во временной области или может быть воплощен так, чтобы включать в себя блок 141 кодирования с возбуждением во временной области и блок 143 кодирования с возбуждением в частотной области. Примеры способов кодирования, применимых к блоку 141 кодирования с возбуждением во временной области, могут включать в себя линейное предсказание с кодовым возбуждением (CELP) или алгебраическое CELP (ACELP), но не ограничиваются этим. Примеры способов кодирования, применимых к блоку 143 кодирования с возбуждением в частотной области, могут включать в себя общее кодирование сигналов (GSC) или кодируемое с преобразованием возбуждение (TCX), но не ограничиваются этим. Может быть эффективным кодировать речевой сигнал в блоке 141 кодирования с возбуждением во временной области, тогда как может быть эффективным кодировать вокальный и/или гармонический сигнал в блоке 143 кодирования с возбуждением в частотной области.

[31] Блок 150 генерирования потока битов может генерировать поток битов так, чтобы он включал в себя режим кодирования, обеспеченный блоком 110 определения режима кодирования, результат кодирования, обеспеченный блоком 130 кодирования в спектральной области, и результат кодирования, обеспеченный блоком 140 кодирования в области линейного предсказания.

[32] Фиг. 2 представляет собой структурную схему, иллюстрирующую конфигурацию устройства 200 аудиокодирования в соответствии с другим примерным вариантом осуществления.

[33] Устройство 200 аудиокодирования, показанное на Фиг. 2, может включать в себя модуль 205 общей предобработки, блок 210 определения режима кодирования, переключающий блок 220, блок 230 кодирования в спектральной области, блок 240 кодирования в области линейного предсказания, а также блок 250 генерирования потока битов. Здесь блок 240 кодирования в области линейного предсказания может включать в себя блок 241 кодирования с возбуждением во временной области и блок 243 кодирования с возбуждением в частотной области, и блок 240 кодирования в области

линейного предсказания, может быть воплощен либо как блок 241 кодирования с возбуждением во временной области, либо как блок 243 кодирования с возбуждением в частотной области. По сравнению с устройством 100 аудиокодирования, показанным на Фиг. 1, устройство 200 аудиокодирования может дополнительно включать в себя модуль 205 общей предобработки, и таким образом описания компонентов, идентичных компонентам устройства 100 аудиокодирования, будут опущены.

[34] Что касается Фиг. 2, модуль 205 общей предобработки может выполнять совместную стереообработку, обработку окружения и/или обработку расширения полосы частот. Совместная стереообработка, обработка окружения и обработка расширения полосы частот могут быть идентичны тем, которые используются некоторым конкретным стандартом, например, стандартом MPEG, но не ограничиваются этим. Вывод модуля 205 общей предобработки может осуществляться в моноканале, стереоканале или мультисканалах. В соответствии с количеством каналов сигнала, выводимого модулем 205 общей предобработки, переключающий блок 220 может включать в себя по меньшей мере один переключатель. Например, если модуль 205 общей предобработки выводит сигнал двух или более каналов, например стереоканала или мультисканала, могут быть обеспечены переключатели, соответствующие этим каналам. Например, первый канал стереосигнала может быть речевым каналом, а второй канал стереосигнала может быть музыкальным каналом. В этом случае аудиосигнал может быть одновременно обеспечен на два переключателя. Дополнительная информация, генерируемая модулем 205 общей предобработки, может быть обеспечена на блок 250 генерирования потока битов и включена в поток битов. Дополнительная информация может быть необходимой для выполнения совместной стереообработки, обработки окружения и/или обработки расширения полосы частот на стороне декодирования, и может включать в себя пространственные параметры, информацию об огибающей, энергетическую информацию и т.д. Однако, также может иметь место различная дополнительная информация, основанная на применяемых технологиях обработки.

[35] В соответствии с одним примерным вариантом осуществления в модуле 205 общей предобработки обработка расширения полосы частот может выполняться по-разному на основе областей кодирования. Аудиосигнал в основной полосе может быть обработан путем использования режима кодирования с возбуждением во временной области или режима кодирования с возбуждением в частотной области, тогда как аудиосигнал в полосе с расширенной полосой частот может обрабатываться во временной области. Обработка расширения полосы частот во временной области может включать в себя множество режимов, в том числе речевой режим или неречевой режим. Альтернативно, аудиосигнал в основной полосе может обрабатываться путем использования режима кодирования в спектральной области, тогда как аудиосигнал в полосе с расширенной полосой частот может обрабатываться в частотной области. Обработка расширения полосы частот в частотной области может включать в себя множество режимов, в том числе переходный режим, нормальный режим или гармонический режим. Для того, чтобы выполнить обработку расширения полосы частот в различных областях, режим кодирования, определенный блоком 110 определения режима кодирования, может быть обеспечен модулю 205 общей предобработки как сигнальная информация. В соответствии с одним примерным вариантом осуществления последняя часть основной полосы и начальная часть полосы с расширенной полосой частот могут до некоторой степени перекрывать друг друга. Положение и размер перекрывающихся частей могут быть установлены заранее.

[36] Фиг. 3 представляет собой структурную схему, иллюстрирующую конфигурацию блока 300 определения режима кодирования в соответствии с одним примерным вариантом осуществления.

5 [37] Блок 300 определения режима кодирования, показанный на Фиг. 3, может включать в себя блок 310 определения первоначального режима кодирования и блок 330 корректировки режима кодирования.

[38] Что касается Фиг. 3, блок 310 определения первоначального режима кодирования может определять, является ли аудиосигнал музыкальным сигналом или речевым сигналом, путем использования параметров особенности, извлекаемых из аудиосигнала.

10 Если аудиосигнал определяется как речевой сигнал, подходящим может быть кодирование в области линейного предсказания. Тем временем, если аудиосигнал определяется как музыкальный сигнал, подходящим может быть кодирование в спектральной области. Блок 310 определения первоначального режима кодирования может определять класс аудиосигнала, указывающий, является ли кодирование в

15 спектральной области, кодирование с возбуждением во временной области или кодирование с возбуждением в частотной области подходящими для аудиосигнала, путем использования параметров особенности, извлекаемых из аудиосигнала. Соответствующий режим кодирования может быть определен на основе класса аудиосигнала. Если переключающий блок (120 на Фиг. 1) имеет две ветви, режим

20 кодирования может выражаться одним битом. Если переключающий блок (120 на Фиг. 1) имеет три ветви, режим кодирования может выражаться двумя битами. Блок 310 определения первоначального режима кодирования может определять, является ли аудиосигнал музыкальным сигналом или речевым сигналом, путем использования

25 любого из различных способов, известных в данной области техники. Примеры таких способов могут включать в себя классификацию FD/LPD или классификацию ACELP/TCX, раскрытые в части кодера стандарта USAC, и классификацию ACELP/TCX, используемую в стандартах AMR, но не ограничиваются этим. Другими словами, первоначальный режим кодирования может быть определен путем использования

30 любого из различных способов, отличных от способа в соответствии с описанными в настоящем документе вариантами осуществления.

[39] Блок 330 корректировки режима кодирования может определять скорректированный режим кодирования путем корректировки первоначального режима кодирования, определенного блоком 310 определения первоначального режима кодирования, с использованием параметров корректировки. В соответствии с одним

35 примерным вариантом осуществления, если в качестве первоначального режима кодирования определяется режим кодирования в спектральной области, первоначальный режим кодирования может быть скорректирован на режим кодирования с возбуждением в частотной области на основе параметров корректировки. Если в качестве первоначального режима кодирования определяется режим кодирования во временной

40 области, первоначальный режим кодирования может быть скорректирован на режим кодирования с возбуждением в частотной области на основе параметров корректировки. Другими словами, определяется, имеется ли ошибка в определении первоначального режима кодирования, путем использования параметров корректировки. Если определено, что ошибки нет в определении первоначального режима кодирования,

45 первоначальный режим кодирования может быть сохранен. С другой стороны, если определено, что имеется ошибка в определении первоначального режима кодирования, первоначальный режим кодирования может быть скорректирован. Корректировка первоначального режима кодирования может быть получена из режима кодирования

в спектральной области на режим кодирования с возбуждением в частотной области и из режима кодирования с возбуждением во временной области на режим кодирования с возбуждением в частотной области.

5 [40] Тем временем, первоначальный режим кодирования или скорректированный режим кодирования могут быть временным режимом кодирования для текущего кадра, где временный режим кодирования для текущего кадра может сравниваться с режимами кодирования для предыдущих кадров в пределах заданной продолжительности затягивания, и может быть определен окончательный режим кодирования для текущего кадра.

10 [41] Фиг. 4 представляет собой структурную схему, иллюстрирующую конфигурацию блока 400 определения первоначального режима кодирования в соответствии с одним примерным вариантом осуществления.

[42] Блок 400 определения первоначального режима кодирования, показанный на Фиг. 4, может включать в себя блок 410 извлечения параметра особенности и блок 430
15 определения.

[43] Со ссылкой на Фиг. 4, блок 410 извлечения параметра особенности может извлекать параметры особенности, необходимые для определения режим кодирования, из аудиосигнала. Примеры извлекаемых параметров особенности включают в себя по
20 меньшей мере один или два из числа параметра основного тона, параметра генерации звука голосовыми связками, параметра корреляции и ошибка линейного предсказания, но не ограничиваются этим. Ниже будут даны подробные описания отдельных параметров.

[44] Во-первых, первый параметр F_1 особенности относится к параметру основного тона, где поведение основного тона может быть определено путем использования N
25 значений основного тона, обнаруживаемых в текущем кадре и по меньшей мере одном предыдущем кадре. Для того, чтобы не допускать эффект от случайного отклонения или неправильного значения основного тона, M значений основного тона, существенно отличающихся от среднего значения N значений основного тона, могут быть удалены. Здесь N и M могут быть значениями, полученными путем экспериментов или
30 предварительных моделирований. Кроме того, N может быть установлено заранее, а разность между значением основного тона, которое должно быть удалено, и средним значением N значений основного тона может быть определена путем экспериментов или предварительных моделирований. Первый параметр F_1 особенности может быть
35 выражен, как показано в Уравнении 1 ниже, путем использования среднего значения $m_{p'}$ и дисперсии $\sigma_{p'}$ в отношении $(N-M)$ значений основного тона.

[45] [Уравнение 1]

$$40 \quad F_1 = \frac{\sigma_{p'}}{m_{p'}}$$

[46] Второй параметр F_2 особенности также относится к параметру основного тона и может указывать достоверность значения основного тона, обнаруживаемого в текущем кадре. Второй параметр F_2 особенности может быть выражен, как показано в Уравнении
45 2 ниже, путем использования дисперсий σ_{SF_1} и σ_{SF_2} значений основного тона, обнаруживаемых соответственно в двух подкадрах SF_1 и SF_2 текущего кадра.

[47] [Уравнение 2]

$$F_2 = \frac{\text{cov}(SF_1, SF_2)}{\sigma_{SF_1} \sigma_{SF_2}}$$

5 [48] Здесь $\text{cov}(SF_1, SF_2)$ обозначает ковариацию между подкадрами SF_1 и SF_2 . Другими словами, второй параметр F_2 особенности указывает корреляцию между двумя подкадрами как расстояние основного тона. В соответствии с одним примерным вариантом осуществления текущий кадр может включать в себя два или более

10 подкадров, и Уравнение 2 может быть модифицировано на основе количества подкадров. [49] Третий параметр F_3 особенности может быть выражен, как показано в Уравнении 3 ниже, на основе параметра генерации звука голосовыми связками, Voicing, и параметра корреляции, Corr.

[50] [Уравнение 3]

$$15 F_3 = \sqrt{Q \frac{c\text{Voicing} - \text{Corr}c^2}{N}}$$

[51] Здесь параметр генерации звука голосовыми связками, Voicing, относится к голосовым особенностям звука и может быть получен любым из различных способов, известных в данной области техники, тогда как параметр корреляции, Corr, может быть

20 получен путем суммирования корреляций между кадрами для каждой полосы. [52] Четвертый параметр F_4 особенности относится к ошибке E_{LPC} линейного предсказания и может быть выражен как показано в Уравнении 4 ниже.

[53] [Уравнение 4]

$$25 F_4 = \frac{\sqrt{(E_{LPCi} - M(E_{LPC}))^2}}{N}$$

[54] Здесь $M(E_{LPC})$ обозначает среднее значение N ошибок линейного предсказания.

30 [55] Блок 430 определения может определять класс аудиосигнала путем использования по меньшей мере одного параметра особенности, обеспечиваемого блоком 410 извлечения параметра особенности, и может определять первоначальный режим кодирования на основе определенного класса. Блок 430 определения может использовать механизм мягкого решения, где по меньшей мере одна смесь может быть сформирована для каждого параметра особенности. В соответствии с одним примерным вариантом осуществления класс аудиосигнала может быть определен путем использования модели смеси распределений Гаусса (GMM) на основе вероятностей смеси. Вероятность $f(x)$ для одной смеси может быть вычислена в соответствии с Уравнением 5, приведенным ниже.

[56] [Уравнение 5]

$$40 f(x) = \frac{1}{\sqrt{(2\pi)^N \det(C^{-1})}} e^{-0.5(x-m)^T C^{-1} (x-m)}$$

$$x = (x_1, \dots, x_N)$$

$$45 m = (cx_1c, \dots, cx_Nc)$$

[57] Здесь x обозначает входной вектор параметра особенности, m обозначает смесь, а c обозначает ковариационную матрицу.

[58] Блок 430 определения может вычислять вероятность P_m музыки и вероятность P_s речи путем использования Уравнения 6, приведенного ниже.

[59] [Уравнение 6]

$$P_m = \prod_{i \in M} p_i, P_s = \prod_{i \in S} p_i$$

[60] Здесь вероятность P_m музыки может быть вычислена путем сложения вероятностей P_i для M смесей, относящихся к параметрам особенности, пригодным для определения музыки, тогда как вероятность P_s речи может быть вычислена путем сложения вероятностей P_i для S смесей, относящихся к параметрам особенности, пригодным для определения речи.

[61] Тем временем для улучшения точности вероятность P_m музыки и вероятность P_s речи могут быть вычислены в соответствии с Уравнением 7, приведенным ниже.

[62] [Уравнение 7]

$$P_m = \prod_{i \in M} p_i(1-p_i^{err}) + \prod_{i \in S} p_i(p_i^{err})$$

$$P_s = \prod_{i \in S} p_i(1-p_i^{err}) + \prod_{i \in M} p_i(p_i^{err})$$

[63] Здесь p_i^{err} обозначает вероятность ошибки каждой смеси. Вероятность ошибки может быть получена путем классификации обучающих данных, включающих в себя чисто речевые сигналы и чисто музыкальные сигналы, используя каждую из смесей и подсчитывая количество неправильных классификаций.

[64] Затем, в соответствии с Уравнением 8, приведенным ниже, могут быть вычислены вероятность P^M того, что все кадры включают в себя только музыкальные сигналы, и вероятность P^S того, что все кадры включают в себя только речевые сигналы, в отношении множества кадров, количество которых равно постоянной продолжительности затягивания. Продолжительность затягивания может быть установлена равной 8, но не ограничивается этим. Восемь кадров могут включать в себя текущий кадр и 7 предыдущих кадров.

[65] [Уравнение 8]

$$P^M = \frac{\prod_{i=0}^{-7} P_m^{(i)}}{\prod_{i=0}^{-7} P_m^{(i)} + \prod_{i=0}^{-7} P_s^{(i)}}$$

$$P^S = \frac{\prod_{i=0}^{-7} P_s^{(i)}}{\prod_{i=0}^{-7} P_m^{(i)} + \prod_{i=0}^{-7} P_s^{(i)}}$$

[66] Далее, множество наборов $\{D_i^M\}$ и $\{D_i^S\}$ условий может быть вычислено с использованием вероятности P_m музыки или вероятности P_s речи, полученных с использованием Уравнения 5 или Уравнения 6. Подробные описания этого будут даны ниже со ссылкой на Фиг. 6. Здесь, можно установить, что каждое условие имеет значение 1 для музыки и имеет значение 0 для речи.

[67] Со ссылкой на Фиг. 6, на этапах 610 и 620 сумма музыкальных условий M и сумма речевых условий S могут быть получены из множества наборов $\{D_i^M\}$ и $\{D_i^S\}$ условий, которые вычисляются путем использования вероятности P_m музыки и вероятности P_s речи. Другими словами, сумма музыкальных условий M и сумма речевых условий S может быть выражена, как показано в Уравнении 9, приведенном ниже.

[68] [Уравнение 9]

$$M = \sum_i Q D_i^M$$

$$S = \sum_i Q D_i^S$$

[69] На этапе 630 сумма музыкальных условий M сравнивается с назначенным пороговым значением T_m . Если сумма музыкальных условий M больше, чем пороговое значение T_m , режим кодирования текущего кадра переключается на режим музыки, то есть на режим кодирования в спектральной области. Если сумма музыкальных условий M меньше или равна пороговому значению T_m , режим кодирования текущего кадра не изменяется.

[70] На этапе 640 сумма речевых условий S сравнивается с назначенным пороговым значением T_s . Если сумма речевых условий S больше, чем пороговое значение T_s , режим кодирования текущего кадра переключается на режим речи, то есть на режим кодирования в области линейного предсказания. Если сумма речевых условий S меньше или равна пороговому значению T_s , режим кодирования текущего кадра не изменяется.

[71] Пороговое значение T_m и пороговое значение T_s могут быть установлены равными значениям, полученным путем экспериментов или предварительного моделирования.

[72] Фиг. 5 представляет собой блок-схему, иллюстрирующую конфигурацию блока 500 извлечения параметров особенности в соответствии с одним примерным вариантом осуществления.

[73] Блок 500 определения первоначального режима кодирования, показанный на Фиг. 5, может включать в себя блок 510 преобразования, блок 520 извлечения спектрального параметра, блок 530 извлечения временного параметра и блок 540 определения.

[74] На Фиг. 5 блок 510 преобразования может преобразовывать первоначальный аудиосигнал из временной области в частотную область. Здесь блок 510 преобразования может применять любой из различных методов преобразования для представления аудиосигнала из временной области в спектральной области. Примеры этих способов могут включать в себя быстрое преобразование Фурье (FFT), дискретное косинусное преобразование (DCT) или модифицированное дискретное косинусное преобразование (MDCT), но не ограничиваются этим.

[75] Блок 520 извлечения спектрального параметра может извлекать по меньшей мере один спектральный параметр из аудиосигнала в частотной области, обеспеченного блоком 510 преобразования. Спектральные параметры могут быть категоризованы на краткосрочные параметры особенности и долгосрочные параметры особенности. Краткосрочные параметры особенности могут быть получены из текущего кадра, тогда как долгосрочные параметры особенности могут быть получены из множества кадров, включающего в себя текущий кадр и по меньшей мере один предыдущий кадр.

[76] Блок 530 извлечения временного параметра может извлекать по меньшей мере один временной параметр из аудиосигнала во временной области. Временные параметры

также могут быть категоризированы на краткосрочные параметры особенности и долгосрочные параметры особенности. Краткосрочные параметры особенности могут быть получены из текущего кадра, тогда как долгосрочные параметры особенности могут быть получены из множества кадров, включающего в себя текущий кадр и по меньшей мере один предыдущий кадр.

[77] Блок определения (430 на Фиг. 4) может определять класс аудиосигнала путем использования спектральных параметров, обеспечиваемых блоком 520 извлечения спектрального параметра, и временных параметров, обеспечиваемых блоком 530 извлечения временного параметра, и может определять первоначальный режим кодирования на основе определенного класса. Блок определения (430 на Фиг. 4) может использовать механизм мягкого решения.

[78] Фиг. 7 представляет собой схему, иллюстрирующую работу блока 310 корректировки режима кодирования в соответствии с одним примерным вариантом осуществления.

[79] Что касается Фиг. 7, на этапе 700 принимается первоначальный режим кодирования, определенный блоком 310 определения первоначального режима кодирования, и может быть определено, является ли режим кодирования режимом во временной области, то есть режимом возбуждения во временной области, или режимом в спектральной области.

[80] На этапе 701, если на этапе 700 определено, что первоначальный режим кодирования является режимом в спектральной области ($state_{TS}=1$), может быть проверен индекс $state_{TTSS}$, указывающий, является ли кодирование с возбуждением в частотной области более подходящим. Индекс $state_{TTSS}$, указывающий, является ли кодирование с возбуждением в частотной области (например, GSC) более подходящим, может быть получен путем использования тональностей различных полос частот. Подробные описания этого будут даны ниже.

[81] Тональность сигнала нижней полосы может быть получена как отношение между суммой множества коэффициентов спектра, имеющих малые значения, в том числе самое малое значение, и коэффициентом спектра, имеющим самое большое значение для данной полосы. Если данными полосами являются 0~1 кГц, 1~2 кГц и 2~4 кГц, тональности t_{01} , t_{12} и t_{24} соответствующих полос и тональность t_L сигнала нижней полосы, то есть основной полосы, могут быть выражены, как показано в Уравнении 10, приведенном ниже.

[82] [Уравнение 10]

$$t_{01} = 0.2 \log_{10} \left(\frac{\max(x_i)}{\prod_{j=0}^{M-1} \text{sort}(x_j)} \right), i, j \in [0, \dots, 1 \text{kHz}]$$

$$t_{12} = 0.2 \log_{10} \left(\frac{\max(x_i)}{\prod_{j=0}^{M-1} \text{sort}(x_j)} \right), i, j \in [1, \dots, 2 \text{kHz}]$$

$$t_{24} = 0.2 \log_{10} \left(\frac{\max(x_i)}{\prod_{j=0}^{M-1} \text{sort}(x_j)} \right), i, j \in [2, \dots, 4 \text{kHz}]$$

$$t_L = \max(t_{01}, t_{12}, t_{24})$$

[83] При этом ошибка линейного предсказания, err , может быть получена путем использования фильтра кодирования с линейным предсказанием (LPC), и может быть использована для того, чтобы удалить сильные тональные компоненты. Другими словами, режим кодирования в спектральной области может быть более эффективным в отношении сильных тональных компонент, чем режим кодирования с возбуждением в частотной области.

[84] Условие начала (“front condition”), $cond_{front}$, для переключения в режим кодирования с возбуждением в частотной области путем использования тональностей и ошибки линейного предсказания, полученных как описано выше, может быть выражено, как показано в Уравнении 11, приведенном ниже.

[85] [Уравнение 11]

$$cond_{front} = t_{12} > t_{12front} \text{ and } t_{24} > t_{24front} \text{ and } t_L > t_{Lfront} \text{ and } err > err_{front}$$

[86] Здесь $t_{12front}$, $t_{24front}$, t_{Lfront} и err_{front} являются пороговыми значениями и могут иметь значения, полученные с помощью экспериментов или предварительных моделирований.

[87] При этом условие конца (“back condition”), $cond_{back}$, для завершения режима кодирования с возбуждением в частотной области путем использования тональностей и ошибки линейного предсказания, полученных как описано выше, может быть выражено, как показано в Уравнении 12, приведенном ниже.

[88] [Уравнение 12]

$$cond_{back} = t_{12} < t_{12back} \text{ and } t_{24} < t_{24back} \text{ and } t_L < t_{Lback}$$

[89] Здесь t_{12back} , t_{24back} , t_{Lback} являются пороговыми значениями и могут иметь значения, полученные с помощью экспериментов или предварительных моделирований.

[90] Другими словами, может быть определено, равен ли единице индекс $state_{TTSS}$, указывающий, является ли кодирование с возбуждением в частотной области (например, GSC) более подходящим, чем кодирование в спектральной области, путем определения, удовлетворяется ли условие начала, показанное в Уравнении 11, или не удовлетворяется ли условие конца, показанное в Уравнении 12. Здесь определение условия конца, показанного в Уравнении 12, может быть необязательным.

[91] На этапе 702, если индекс $state_{TTSS}$ равен 1, режим кодирования с возбуждением

в частотной области может быть определен как окончательный режим кодирования. В этом случае режим кодирования в спектральной области, который является первоначальным режимом кодирования, корректируется на режим кодирования с возбуждением в частотной области, который является окончательным режимом кодирования.

[92] На этапе 705, если на этапе 701 определено, что индекс $state_{TTSS}$ равен 0, может быть проверен индекс $state_{SS}$ для того, чтобы определить, включает ли в себя аудиосигнал сильную речевую характеристику. Если имеется ошибка в определении режима кодирования в спектральной области, режим кодирования с возбуждением в частотной области может быть более эффективным, чем режим кодирования в спектральной области. Индекс $state_{SS}$ для определения того, включает ли в себя аудиосигнал сильную речевую характеристику, может быть получен путем использования разности vc между параметром генерации звука голосовыми связками и параметром корреляции.

[93] Условие начала, $cond_{front}$, для переключения на режим сильной речи путем использования разности vc между параметром генерации звука голосовыми связками и параметром корреляции может быть выражено, как показано в Уравнении 13, приведенном ниже.

[94] [Уравнение 13]

$$cond_{front} = vc > vc_{front}$$

[95] Здесь vc_{front} представляет собой пороговое значение и может иметь значение, полученное с помощью экспериментов или предварительных моделирований.

[96] При этом условие конца, $cond_{back}$, для завершения режима сильной речи путем использования разности vc между параметром генерации звука голосовыми связками и параметром корреляции может быть выражено, как показано в Уравнении 14, приведенном ниже.

[97] [Уравнение 14]

$$cond_{back} = vc < vc_{back}$$

[98] Здесь vc_{back} представляет собой пороговое значение и может иметь значение, полученное с помощью экспериментов или предварительных моделирований.

[99] Другими словами, на этапе 705 может быть определено, равен ли единице индекс $state_{SS}$, указывающий, является ли кодирование с возбуждением в частотной области (например, GSC) более подходящим, чем кодирование в спектральной области, путем определения, удовлетворяется ли условие начала, показанное в Уравнении 13, или не удовлетворяется ли условие конца, показанное в Уравнении 14. Здесь определение условия конца, показанного в Уравнении 14, может быть необязательным.

[100] На этапе 706, если на этапе 705 определено, что индекс $state_{SS}$ равен 0, то есть аудиосигнал не включает в себя сильной речевой характеристики, режим кодирования в спектральной области может быть определен как окончательный режим кодирования. В этом случае режим кодирования в спектральной области, который является первоначальным режимом кодирования, сохраняется как окончательный режим кодирования.

[101] На этапе 707, если на этапе 705 определено, что индекс $state_{SS}$ равен 1, то есть аудиосигнал включает в себя сильную речевую характеристику, режим кодирования с

возбуждением в частотной области может быть определен как окончательный режим кодирования. В этом случае режим кодирования в спектральной области, который является первоначальным режимом кодирования, корректируется на режим кодирования с возбуждением в частотной области, который является окончательным режимом кодирования.

[102] Путем выполнения этапов 700, 701 и 705 ошибка в определении режима кодирования в спектральной области в качестве первоначального режима кодирования может быть скорректирована. Более подробно, режим кодирования в спектральной области, который является первоначальным режимом кодирования, может быть сохранен или переключен на режим кодирования с возбуждением в частотной области в качестве окончательного режима кодирования.

[103] При этом, если на этапе 700 определено, что первоначальный режим кодирования является режимом кодирования в области линейного предсказания ($state_{TS}=0$), может быть проверен индекс $state_{SM}$ для определения того, включает ли в себя аудиосигнал сильную музыкальную характеристику. Если имеется ошибка в определении режима кодирования в области линейного предсказания, то есть режима кодирования с возбуждением во временной области, режим кодирования с возбуждением в частотной области может быть более эффективным, чем режим кодирования с возбуждением во временной области. $state_{SM}$ для определения того, включает ли в себя аудиосигнал сильную музыкальную характеристику, может быть получен путем использования значения $1-vc$, получаемого вычитанием из единицы разности vc между параметром генерации звука голосовыми связками и параметром корреляции.

[104] Условие начала, $cond_{front}$, для переключения на режим сильной музыки путем использования значения $1-vc$, получаемого вычитанием из единицы разности vc между параметром генерации звука голосовыми связками и параметром корреляции, может быть выражено, как показано в Уравнении 15, приведенном ниже.

[105] [Уравнение 15]

$$cond_{front} = 1 - vc > vcm_{front}$$

[106] Здесь vcm_{front} представляет собой пороговое значение и может иметь значение, получаемое посредством экспериментов или предварительных моделирований.

[107] При этом условие конца $cond_{back}$ для завершения режима сильной музыки путем использования значения $1-vc$, получаемого вычитанием из единицы разности vc между параметром генерации звука голосовыми связками и параметром корреляции, может быть выражено, как показано в Уравнении 16, приведенном ниже.

[108] [Уравнение 16]

$$cond_{back} = 1 - vc < vcm_{back}$$

[109] Здесь vcm_{back} представляет собой пороговое значение и может иметь значение, полученное с помощью экспериментов или предварительных моделирований.

[110] Другими словами, на этапе 709 может быть определено, равен ли единице индекс $state_{SM}$, указывающий, является ли кодирование с возбуждением в частотной области (например, GSC) более подходящим, чем кодирование с возбуждением во временной области, путем определения, удовлетворяется ли условие начала, показанное в Уравнении 15, или не удовлетворяется ли условие конца, показанное в Уравнении 16. Здесь определение условия конца, показанного в Уравнении 16, может быть

необязательным.

[111] На этапе 710, если на этапе 709 определено, что индекс $state_{SM}$ равен 0, то есть аудиосигнал не включает в себя сильной музыкальной характеристики, режим кодирования с возбуждением во временной области может быть определен как окончательный режим кодирования. В этом случае режим кодирования в области линейного предсказания, который является первоначальным режимом кодирования, переключается на режим кодирования с возбуждением во временной области в качестве окончательного режима кодирования. В соответствии с одним примерным вариантом осуществления можно считать, что первоначальный режим кодирования сохраняется без изменений, если режим кодирования в области линейного предсказания соответствует режиму кодирования с возбуждением во временной области.

[112] На этапе 707, если на этапе 709 определено, что индекс $state_{SM}$ равен 1, то есть аудиосигнал включает в себя сильную музыкальную характеристику, режим кодирования с возбуждением в частотной области может быть определен как окончательный режим кодирования. В этом случае режим кодирования в области линейного предсказания, который является первоначальным режимом кодирования, корректируется на режим кодирования с возбуждением в частотной области, который является окончательным режимом кодирования.

[113] Ошибка в определении первоначального режима кодирования может быть скорректирована путем выполнения этапов 700 и 709. Более подробно, режим кодирования в области линейного предсказания (например, режим кодирования с возбуждением во временной области), который является первоначальным режимом кодирования, может быть сохранен или переключен на режим кодирования с возбуждением в частотной области в качестве окончательного режима кодирования.

[114] В соответствии с одним примерным вариантом осуществления этап 709 для определения, включает ли в себя аудиосигнал сильную музыкальную характеристику, для корректировки ошибки в определении режима кодирования в области линейного предсказания, может быть необязательным.

[115] В соответствии с другим примерным вариантом осуществления последовательность выполнения этапа 705 для определения, включает ли в себя аудиосигнал сильную речевую характеристику, и этапа 701 для определения, является ли режим кодирования с возбуждением в частотной области подходящим, может быть изменена на обратную. Другими словами, после выполнения этапа 700 сначала может быть выполнен этап 705, а затем может быть выполнен этап 701. В этом случае параметры, используемые для определений, могут быть изменены в соответствии с необходимостью.

[116] Фиг. 8 представляет собой структурную схему, иллюстрирующую конфигурацию устройства 800 аудиодекодирования в соответствии с одним примерным вариантом осуществления.

[117] Устройство 800 аудиодекодирования, показанное на Фиг. 8, может включать в себя блок 810 анализа потока битов, блок 820 декодирования в спектральной области, блок 830 декодирования в области линейного предсказания и переключающий блок 840. Блок 830 декодирования в области линейного предсказания может включать в себя блок 831 декодирования с возбуждением во временной области и блок 833 декодирования с возбуждением в частотной области, где блок 830 декодирования в области линейного предсказания может быть воплощен как по меньшей мере один из блока 831 декодирования с возбуждением во временной области и блока 833 декодирования с возбуждением в частотной области. Если нет необходимости

воплощения в виде отдельного аппаратного обеспечения, вышеперечисленные компоненты могут быть интегрированы в по меньшей мере один модуль и могут быть осуществлены как по меньшей мере один процессор (не показан).

5 [118] Со ссылкой Фиг. 8, блок 810 анализа потока битов может анализировать принятый поток битов и разделять информацию о режиме кодирования и кодируемые данные. Режим кодирования может соответствовать либо первоначальному режиму кодирования, полученному путем определения одного из множества режимов кодирования, включающего в себя первый режим кодирования и второй режим кодирования в соответствии с характеристиками аудиосигнала, либо третьему режиму кодирования, скорректированному из первоначального режима кодирования, если
10 имеется ошибка в определении первоначального режима кодирования.

[119] Блок 820 декодирования в спектральной области может декодировать данные, закодированные в спектральной области, из разделенных закодированных данных.

15 [120] Блок 830 декодирования в области линейного предсказания может декодировать данные, закодированные в области линейного предсказания, из разделенных закодированных данных. Если блок 830 декодирования в области линейного предсказания включает в себя блок 831 декодирования с возбуждением во временной области и блок 833 декодирования с возбуждением в частотной области, блок 830 декодирования в области линейного предсказания может выполнять в отношении
20 разделенных закодированных данных декодирование с возбуждением во временной области или декодирование с возбуждением в частотной области.

[121] Переключающий блок 840 может переключать либо сигнал, реконструированный блоком 820 декодирования в спектральной области, либо сигнал, реконструированный блоком 830 декодирования в области линейного предсказания,
25 и может обеспечивать переключенный сигнал в качестве конечного реконструированного сигнала.

[122] Фиг. 9 представляет собой структурную схему, иллюстрирующую конфигурацию устройства 900 аудиодекодирования в соответствии с другим примерным вариантом осуществления.

30 [123] Устройство 900 аудиодекодирования может включать в себя блок 910 анализа потока битов, блок 920 декодирования в спектральной области, блок 930 декодирования в области линейного предсказания, переключающий блок 940 и модуль 950 общей постобработки. Блок 930 декодирования в области линейного предсказания может включать в себя блок 931 декодирования с возбуждением во временной области и блок
35 933 декодирования с возбуждением в частотной области, где блок 930 декодирования в области линейного предсказания может быть воплощен как по меньшей мере один из блока 931 декодирования с возбуждением во временной области и блока 933 декодирования с возбуждением в частотной области. Если нет необходимости воплощения в виде отдельного аппаратного обеспечения, вышеперечисленные
40 компоненты могут быть интегрированы в по меньшей мере один модуль и могут быть осуществлены как по меньшей мере один процессор (не показан). По сравнению с устройством 800 аудиодекодирования, показанным на Фиг. 8, устройство 900 аудиодекодирования может дополнительно включать в себя модуль 950 общей постобработки, и таким образом описания компонентов, идентичных компонентам
45 устройства 800 аудиодекодирования, будут опущены.

[124] Со ссылкой Фиг. 9, модуль 950 общей постобработки может выполнять совместную стереообработку, обработку окружения и/или обработку расширения полосы частот в соответствии с модулем общей предобработки (205 на Фиг. 2).

[125] Способы в соответствии с примерными вариантами осуществления могут быть записаны как исполняемые компьютером программы и могут быть реализованы в цифровых компьютерах общего назначения, которые исполняют программы путем использования невременного считываемого компьютером носителя записи. В

5 дополнение к этому структуры данных, программные инструкции или файлы данных, которые могут использоваться в этих вариантах осуществления, могут быть записаны на невременном считываемом компьютером носителе записи различными способами. Невременной считываемый компьютером носитель записи представляет собой любое устройство хранения данных, которое может хранить данные, которые могут быть

10 затем считаны компьютерной системой. Примеры невременного считываемого компьютером носителя записи включают в себя магнитные носители записи, такие как жесткие диски, дискеты и магнитные ленты, оптические носители записи, такие как CD-ROM и DVD, магнитооптические носители, такие как оптические диски, а также аппаратные устройства, такие как ROM, RAM и флэш-память, специально

15 конфигурируемые для того, чтобы хранить и исполнять программные инструкции. В дополнение к этому, невременной считываемый компьютером носитель записи может быть передающей средой для передачи сигнала, назначающего программные инструкции, структуры данных и т.п. Примеры программных инструкций могут включать в себя не только коды механического языка, создаваемые компилятором, но

20 также и коды языка высокого уровня, исполняемые компьютером с использованием интерпретатора и т.п.

[126] В то время как примерные варианты осуществления были конкретно показаны и описаны выше, специалисту в данной области техники будет понятно, что в них могут быть произведены различные изменения в форме и деталях без отступления от сущности

25 и объема концепции настоящего изобретения, определяемого прилагаемой формулой изобретения. Примерные варианты осуществления следует рассматривать только в описательном смысле, а не в целях ограничения. Следовательно, объем концепции настоящего изобретения определяется не подробным описанием примерных вариантов осуществления, а прилагаемой формулой изобретения, и все различия, находящиеся в

30 объеме, должны рассматриваться как включенные в представленную концепцию настоящего изобретения.

(57) Формула изобретения

1. Способ определения режима кодирования, содержащий:
 - 35 определение класса текущего кадра в качестве одного из музыкального класса и речевого класса на основе характеристик сигнала;
 - получение параметров особенностей, включающих в себя соответствующие тональности в по меньшей мере двух частотных областях и ошибку линейного предсказания, на основе множества кадров, включающего в себя упомянутый текущий
 - 40 кадр;
 - определение, возникает ли ошибка в определенном классе текущего кадра, на основе параметров особенностей;
 - когда возникает ошибка в определенном классе текущего кадра и определенным классом текущего кадра является музыкальный класс, корректировку определенного
 - 45 класса текущего кадра в речевой класс; и
 - когда возникает ошибка в определенном классе текущего кадра и определенным классом текущего кадра является речевой класс, корректировку определенного класса текущего кадра в музыкальный класс.

2. Способ по п. 1, в котором параметры особенностей дополнительно содержат разность между параметром генерации звука голосовыми связками и параметром корреляции.

5

10

15

20

25

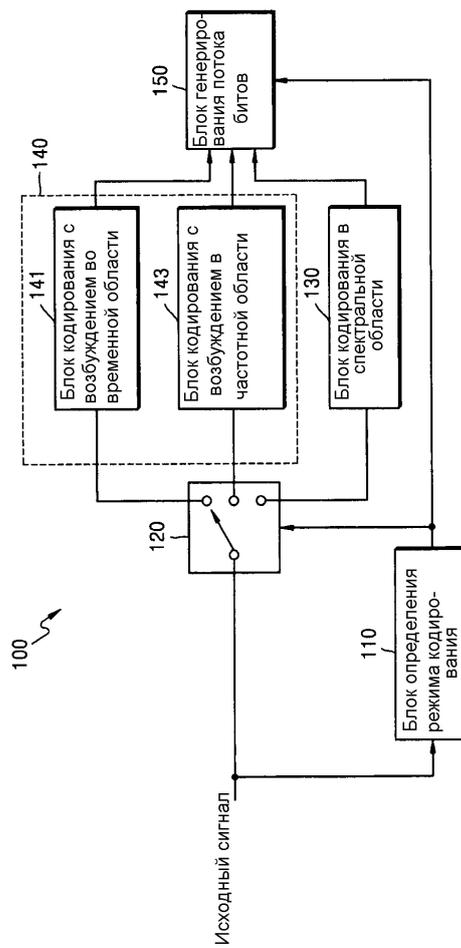
30

35

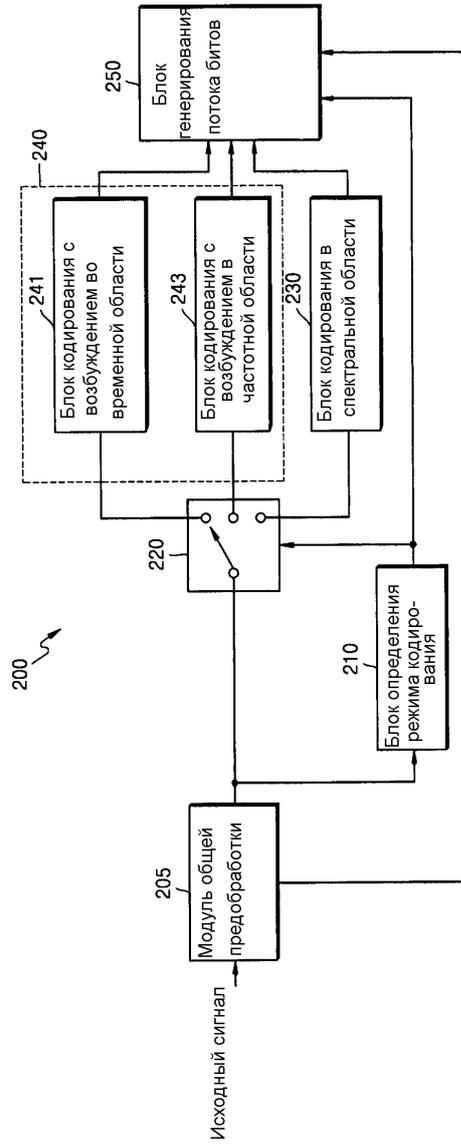
40

45

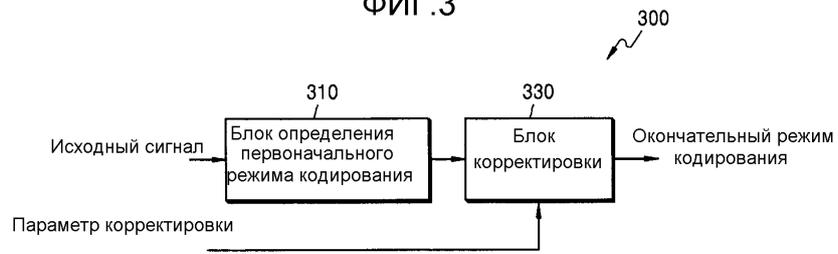
ФИГ.1



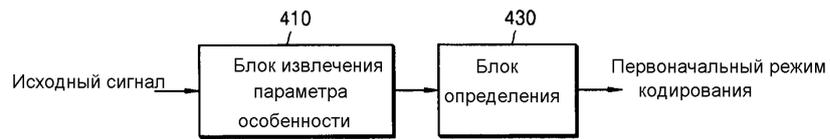
ФИГ.2



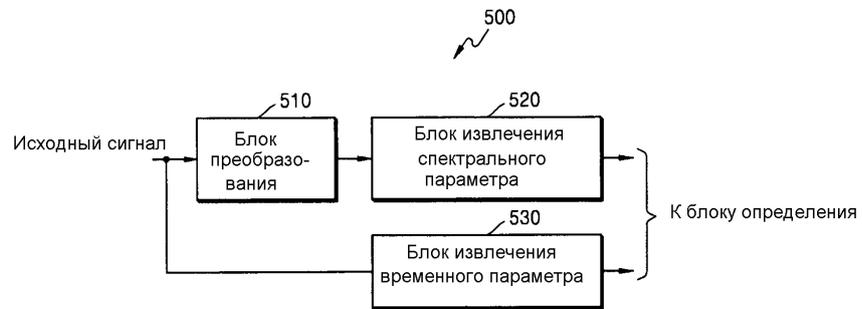
ФИГ.3



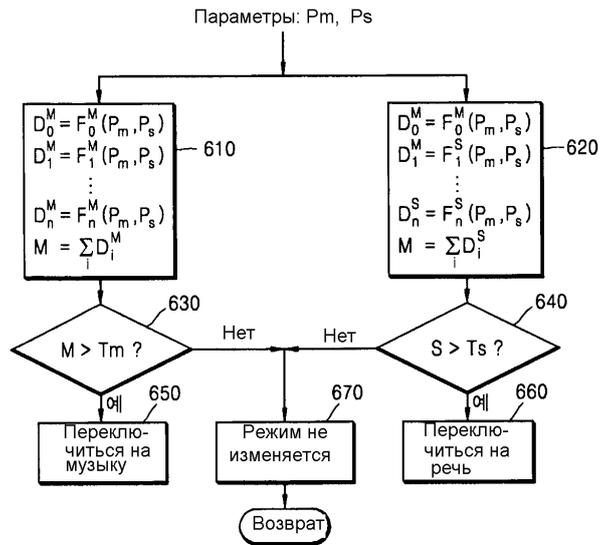
ФИГ.4



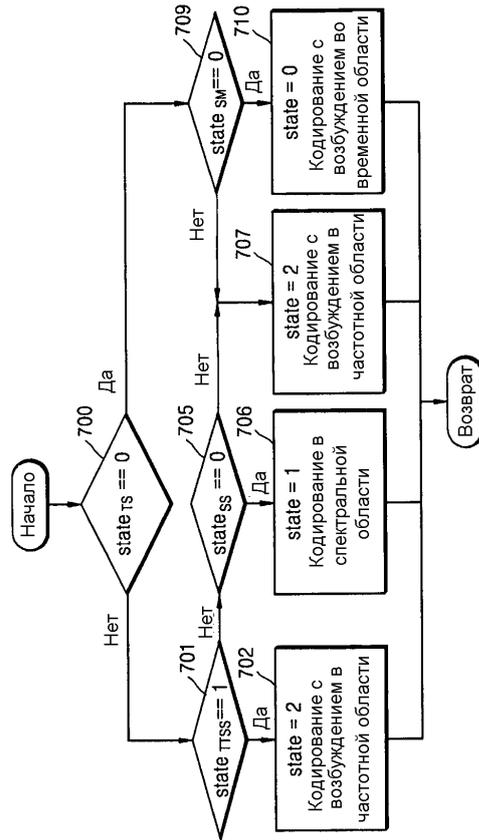
ФИГ.5



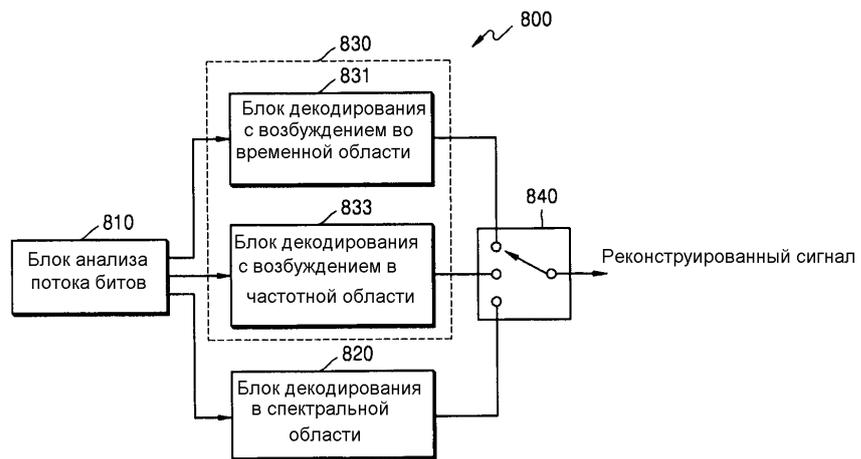
ФИГ.6



ФИГ.7



ФИГ.8



ФИГ.9

