



ФЕДЕРАЛЬНАЯ СЛУЖБА
ПО ИНТЕЛЛЕКТУАЛЬНОЙ СОБСТВЕННОСТИ

(12) ОПИСАНИЕ ИЗОБРЕТЕНИЯ К ПАТЕНТУ

(21)(22) Заявка: 2015149810, 21.07.2015

(24) Дата начала отсчета срока действия патента:
21.07.2015

Дата регистрации:
02.10.2017

Приоритет(ы):

(30) Конвенционный приоритет:
28.07.2014 EP 14178809.1

(43) Дата публикации заявки: 23.05.2017 Бюл. № 15

(45) Опубликовано: 02.10.2017 Бюл. № 28

(85) Дата начала рассмотрения заявки РСТ на
национальной фазе: 20.11.2015

(86) Заявка РСТ:
EP 2015/066677 (21.07.2015)

(87) Публикация заявки РСТ:
WO 2016/016053 (04.02.2016)

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

(72) Автор(ы):

РАВЕЛЛИ, Эммануэль (DE),
МУЛЬТРУС, Маркус (DE),
ДЕЛА, Штефан (DE),
ГРИЛЛ, Бернхард (DE),
ЯНДЕР, Мануэль (DE)

(73) Патентообладатель(и):

ФРАУНХОФЕР-ГЕЗЕЛЛЬШАФТ ЦУР
ФЕРДЕРУНГ ДЕР АНГЕВАНДТЕН
ФОРШУНГ Е.Ф. (DE)

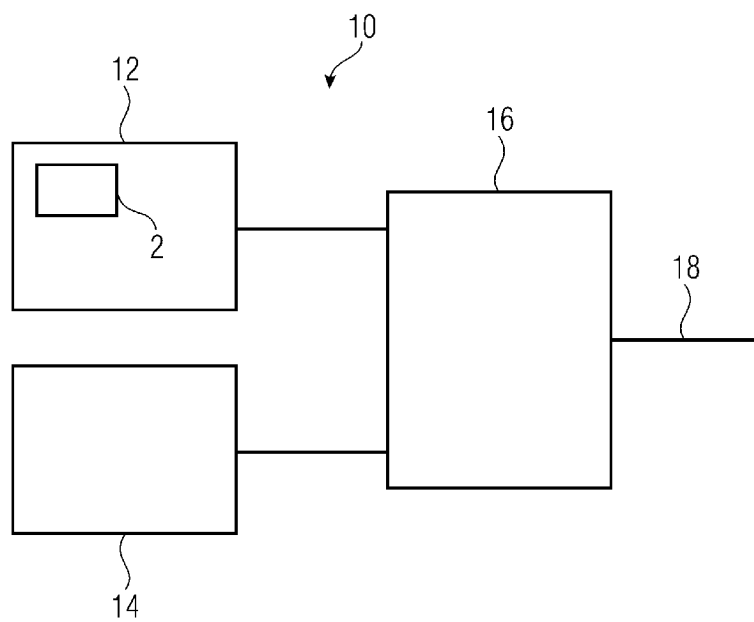
(56) Список документов, цитированных в отчете
о поиске: WO 2012/110448 A1, 23.08.2012. US
2008/0312914 A1, 18.12.2008. US 2013/0332177
A1, 12.12.2013. RU 2483366 C2, 27.05.2013. RU
2439721 C2, 10.01.2012. WO 2007/051548 A1,
10.05.2007. US 2009/0012797 A1, 08.01.2009. EP
1396843 B1, 15.05.2013.

(54) УСТРОЙСТВО И СПОСОБ ВЫБОРА ОДНОГО ИЗ ПЕРВОГО АЛГОРИТМА КОДИРОВАНИЯ И ВТОРОГО АЛГОРИТМА КОДИРОВАНИЯ С ИСПОЛЬЗОВАНИЕМ УМЕНЬШЕНИЯ ГАРМОНИК

(57) Реферат:

Изобретение относится к средствам для выбора алгоритма кодирования. Технический результат заключается в уменьшении сложности выбора между первым алгоритмом кодирования и вторым алгоритмом кодирования. Устройство для выбора одного из первого алгоритма кодирования и второго алгоритма кодирования, для кодирования части аудиосигнала, чтобы получать кодированную версию части аудиосигнала, содержит фильтр, выполненный с возможностью принимать аудиосигнал, уменьшать амплитуду гармоник в аудиосигнале и выводить фильтрованную версию аудиосигнала. Первый модуль оценки предоставляется для

использования фильтрованной версии аудиосигнала при оценке SNR или сегментального SNR части аудиосигнала в качестве первого показателя качества для части аудиосигнала, которая ассоциирована с первым алгоритмом кодирования, без фактического кодирования и декодирования части аудиосигнала с использованием первого алгоритма кодирования. Второй модуль оценки предоставляется для оценки SNR или сегментального SNR в качестве второго показателя качества для части аудиосигнала, которая ассоциирована со вторым алгоритмом кодирования. 5 н. и 10 з.п. ф-лы, 5 ил.



ФИГ. 1



FEDERAL SERVICE
FOR INTELLECTUAL PROPERTY

(51) Int. Cl.

G10L 19/22 (2013.01)*G10L 19/02* (2013.01)*G10L 19/08* (2013.01)*G10L 19/09* (2013.01)**(12) ABSTRACT OF INVENTION**(21)(22) Application: **2015149810, 21.07.2015**(24) Effective date for property rights:
21.07.2015Registration date:
02.10.2017

Priority:

(30) Convention priority:
28.07.2014 EP 14178809.1(43) Application published: **23.05.2017** Bull. № 15(45) Date of publication: **02.10.2017** Bull. № 28(85) Commencement of national phase: **20.11.2015**(86) PCT application:
EP 2015/066677 (21.07.2015)(87) PCT publication:
WO 2016/016053 (04.02.2016)Mail address:
**129090, Moskva, ul. B.Spasskaya, 25, stroenie 3,
OOO "Yuridicheskaya firma Gorodisskij i Partnery"**

(72) Inventor(s):

**RAVELLI, Emmanuel (DE),
MULTRUS, Markus (DE),
DELA, Shtefan (DE),
GRILL, Bernkhard (DE),
YANDER, Manuel (DE)**

(73) Proprietor(s):

**FRAUNKHOFER-GEZELLSCHAFT TSUR
FERDERUNG DER ANGEVANDTEN
FORSHUNG E.F. (DE)****(54) DEVICE AND METHOD OF SELECTION OF ONE OF FIRST CODING ALGORITHM AND SECOND CODING ALGORITHM BY USING HARMONIC REDUCTION**

(57) Abstract:

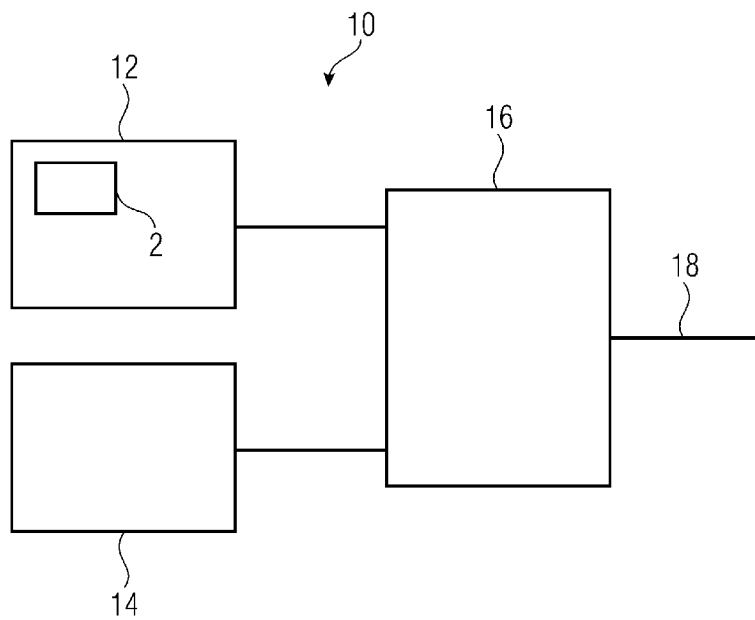
FIELD: physics.

SUBSTANCE: device for selecting one of the first coding algorithm and the second coding algorithm to encode a portion of the audio signal to obtain an encoded version of a portion of the audio signal comprises a filter configured to receive an audio signal, reduce the amplitude of harmonics in the audio signal, and output a filtered version of the audio signal. The first evaluation module is provided to use the filtered version of the audio signal when estimating an SNR or segmental SNR portion of an audio signal as a first quality score for a portion of an audio signal that is

associated with the first coding algorithm without actually encoding and decoding a portion of the audio signal using the first coding algorithm. The second evaluation module is provided for estimating the SNR or segment SNR as the second quality score for the portion of the audio signal that is associated with the second encoding algorithm.

EFFECT: reducing the complexity of the choice between the first coding algorithm and the second coding algorithm.

15 cl, 5 dwg



ФИГ. 1

Подробное описание изобретения

Настоящее изобретение относится к кодированию аудио и, в частности, к переключаемому кодированию аудио, при котором для различных частей аудиосигнала, кодированный сигнал формируется с использованием различных алгоритмов

5 кодирования.

Известны переключаемые аудиокодеры, которые определяют различные алгоритмы кодирования для различных частей аудиосигнала. Обычно, переключаемые аудиокодеры обеспечивают переключение между двумя различными режимами, т.е. алгоритмами, такими как ACELP (линейное прогнозирование с возбуждением по алгебраическому

10 коду) и TCX (возбуждение по кодированию с преобразованием).

LPD-режим MPEG USAC (стандартизированного кодирования речи и аудио по MPEG-стандарту) основан на двух различных режимах ACELP и TCX. ACELP предоставляет лучшее качество для речевых и переходных сигналов. TCX предоставляет лучшее качество для музыкальных и шумоподобных сигналов. Кодер определяет то, какой

15 режим следует использовать, на покадровой основе. Решение, принимаемое посредством кодера, является критически важным для качества кодека. Одно неправильное решение может вызывать сильный артефакт, в частности, на низких скоростях передачи битов. Наиболее простой подход для определения того, какой режим использовать, представляет собой выбор режима с замкнутым контуром, т.е. выполнение полного

20 кодирования/декодирования обоих режимов, затем вычисление критериев выбора (например, сегментального SNR) для обоих режимов на основе аудиосигнала и кодированных/декодированных аудиосигналов и в завершение выбор режима на основе критериев выбора. Этот подход, в общем, обеспечивает стабильное и надежное решение. Тем не менее, он также требует большого объема сложности, поскольку оба режима

25 должны выполняться в каждом кадре. Чтобы уменьшать сложность, альтернативный подход представляет собой выбор режима с разомкнутым контуром. Выбор с разомкнутым контуром состоит не из выполнения полного кодирования/декодирования обоих режимов, а вместо этого выбора одного режима с использованием критериев выбора, вычисленных с низкой

30 сложностью. Сложность по принципу наихудшего случая затем уменьшается посредством сложности наименее сложного режима (обычно TCX), минус сложность, требуемая для того, чтобы вычислять критерии выбора. Снижение сложности обычно является значительным, что делает этот вид подхода практически полезным, когда сложность кодека по принципу наихудшего случая ограничивается.

35 AMR-WB+-стандарт (заданный в Международном стандарте 3GPP TS 26.290 V6.1.0 2004-12) включает в себя выбор режима с разомкнутым контуром, используемый для того, чтобы принимать решение с учетом всех комбинаций ACELP/TCX20/TCX40/TCX80 в кадре в 80 мс. Он описывается в разделе 5.2.4 3GPP TS 26.290. Он также описывается в докладе на конференции "Low Complex Audio Encoding for Mobile, Multimedia", VTC

40 2006, авторов Makinen и др., и в US 7747430 B2 и US7739120 B2, принадлежащих автору этого доклада на конференции.

US 7747430 B2 раскрывает выбор режима с разомкнутым контуром на основе анализа параметров долгосрочного прогнозирования. US 7739120 B2 раскрывает выбор режима с разомкнутым контуром на основе характеристик сигналов, указывающих

45 тип аудиоконтента в соответствующих секциях аудиосигнала, при этом если такой выбор не является практически осуществимым, выбор дополнительно основан на статистической оценке, выполняемой для соответствующих соседних секций.

Выбор режима с разомкнутым контуром AMR-WB+ может описываться на двух

основных этапах. На первом основном этапе, для аудиосигнала вычисляются несколько признаков, такие как среднеквадратическое отклонение энергетических уровней, отношение низкочастотной/высокочастотной энергии, полная энергия, расстояние между ISP (парами спектральных иммитансов), запаздывания и усиления основного тона, спектральный наклон. Эти признаки затем используются для того, чтобы проводить выбор между ACELP и TCX, с использованием простого классификатора на основе порогового значения. Если TCX выбирается на первом основном этапе, то второй основной этап принимает решение с учетом возможных комбинаций TCX20/TCX40/TCX80 в режиме с замкнутым контуром.

WO 2012/110448 A1 раскрывает подход для принятия решения с учетом двух алгоритмов кодирования, имеющих различные характеристики, на основе результата обнаружения переходных частей и результата в отношении качества аудиосигнала. Помимо этого, раскрыто применение гистерезиса, при этом гистерезис основывается на операциях выбора, выполняемых в прошлом, т.е. для более ранних частей аудиосигнала.

В докладе на конференции "Low Complex Audio Encoding for Mobile, Multimedia", VTC 2006, авторов Makinen и др., выбор режима с замкнутым контуром и с разомкнутым контуром AMR-WB+ сравнивается. Субъективные тесты на основе прослушивания указывают то, что выбор режима с разомкнутым контуром работает значительно хуже по сравнению с выбором режима с замкнутым контуром. Тем не менее, также показано, что выбор режима с разомкнутым контуром уменьшает сложность по принципу наихудшего случая на 40%.

Цель изобретения заключается в том, чтобы предоставлять улучшенный подход, который обеспечивает выбор между первым алгоритмом кодирования и вторым алгоритмом кодирования с хорошей производительностью и меньшей сложностью.

Эта цель достигается посредством устройства по п. 1, способа по п. 18 и компьютерной программы по п. 19.

Варианты осуществления изобретения предоставляют устройство для выбора одного из первого алгоритма кодирования, имеющего первую характеристику, и второго алгоритма кодирования, имеющего вторую характеристику, для кодирования части аудиосигнала, чтобы получать кодированную версию части аудиосигнала, содержащее:

- фильтр, выполненный с возможностью принимать аудиосигнал, уменьшать амплитуду гармоник в аудиосигнале и выводить фильтрованную версию аудиосигнала;

- первый модуль оценки для использования фильтрованной версии аудиосигнала при оценке SNR (отношения "сигнал-шум") или сегментированного SNR части аудиосигнала в качестве первого показателя качества для части аудиосигнала, которая ассоциирована с первым алгоритмом кодирования, без фактического кодирования и декодирования части аудиосигнала с использованием первого алгоритма кодирования;

- второй модуль оценки для оценки SNR или сегментированного SNR в качестве второго показателя качества для части аудиосигнала, которая ассоциирована со вторым алгоритмом кодирования, без фактического кодирования и декодирования части аудиосигнала с использованием второго алгоритма кодирования; и

- контроллер для выбора первого алгоритма кодирования или второго алгоритма кодирования на основе сравнения между первым показателем качества и вторым показателем качества.

Варианты осуществления изобретения предоставляют способ выбора одного из первого алгоритма кодирования, имеющего первую характеристику, и второго алгоритма кодирования, имеющего вторую характеристику, для кодирования части

аудиосигнала, чтобы получать кодированную версию части аудиосигнала, содержащий:

- фильтрацию аудиосигнала, чтобы уменьшать амплитуду гармоник в аудиосигнале и выводить фильтрованную версию аудиосигнала;

- использование фильтрованной версии аудиосигнала при оценке SNR или сегментального SNR части аудиосигнала в качестве первого показателя качества для части аудиосигнала, которая ассоциирована с первым алгоритмом кодирования, без фактического кодирования и декодирования части аудиосигнала с использованием первого алгоритма кодирования;

- оценку второго показателя качества для части аудиосигнала, которая ассоциирована со вторым алгоритмом кодирования, без фактического кодирования и декодирования части аудиосигнала с использованием второго алгоритма кодирования; и

- выбор первого алгоритма кодирования или второго алгоритма кодирования на основе сравнения между первым показателем качества и вторым показателем качества.

Варианты осуществления изобретения основаны на признании того факта, что выбор с разомкнутым контуром с повышенной производительностью может реализовываться посредством оценки показателя качества для каждого из первого и второго алгоритмов кодирования и выбора одного из алгоритмов кодирования на основе сравнения между первым и вторым показателями качества. Показатели качества оцениваются, т.е.

аудиосигнал фактически не кодируется и декодируется, чтобы получать показатели качества. Таким образом, показатели качества могут получаться с меньшей сложностью. Выбор режима затем может выполняться с использованием оцененных показателей качества, сравнимых с выбором режима с замкнутым контуром. Кроме того, изобретение основано на признании того факта, что улучшенный выбор режима может получаться, если оценка первого показателя качества использует фильтрованную версию части аудиосигнала, в которой гармоники уменьшаются по сравнению с нефильтовой версией аудиосигнала.

В вариантах осуществления изобретения, реализуется выбор режима с разомкнутым контуром, при котором сегментальное SNR ACELP и TCX сначала оценивается с низкой сложностью. Кроме того, затем выбор режима выполняется с использованием этих значений оцененного сегментального SNR, как при выборе режима с замкнутым контуром.

Варианты осуществления изобретения не используют классический подход "признаки+классификатор", который осуществляется при выборе режима с разомкнутым контуром в AMR-WB+. Тем не менее, вместо этого, варианты осуществления изобретения пытаются оценивать показатель качества каждого режима и выбирать режим, который обеспечивает наилучшее качество.

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

Фиг. 1 показывает схематичный вид варианта осуществления устройства для выбора одного из первого алгоритма кодирования и второго алгоритма кодирования;

Фиг. 2 показывает схематичный вид варианта осуществления устройства для кодирования аудиосигнала;

Фиг. 3 показывает схематичный вид варианта осуществления устройства для выбора одного из первого алгоритма кодирования и второго алгоритма кодирования;

Фиг. 4а и 4б являются возможными представлениями SNR и сегментального SNR.

В нижеприведенном описании, аналогичные элементы/этапы на различных чертежах обозначаются посредством идентичных ссылок с номерами. Следует отметить, что на чертежах опущены такие признаки, как сигнальные соединения и т.п., которые не

требуются при понимании изобретения.

Фиг. 1 показывает устройство 10 для выбора одного из первого алгоритма кодирования, такого как ТСХ-алгоритм, и второго алгоритма кодирования, такого как ACELP-алгоритм, в качестве кодера для кодирования части аудиосигнала. Устройство 10 содержит первый модуль 12 оценки для оценки SNR или сегментального SNR части аудиосигнала в качестве первого показателя качества для части сигнала предоставляется. Первый показатель качества ассоциирован с первым алгоритмом кодирования. Устройство 10 содержит фильтр 2, выполненный с возможностью принимать аудиосигнал, уменьшать амплитуду гармоник в аудиосигнале и выводить 10 фильтрованную версию аудиосигнала. Фильтр 2 может быть внутренним для первого модуля 12 оценки, как показано на фиг. 1, или может быть внешним для первого модуля 12 оценки. Первый модуль 12 оценки использует фильтрованную версию аудиосигнала при оценке первого показателя качества. Другими словами, первый модуль 12 оценки оценивает первый показатель качества, который должна иметь часть аудиосигнала при 15 кодировании и декодировании с использованием первого алгоритма кодирования, без фактического кодирования и декодирования части аудиосигнала с использованием первого алгоритма кодирования. Устройство 10 содержит второй модуль 14 оценки для оценки второго показателя качества для части сигнала. Второй показатель качества ассоциирован со вторым алгоритмом кодирования. Другими словами, второй модуль 20 14 оценки оценивает второй показатель качества, который должна иметь часть аудиосигнала при кодировании и декодировании с использованием второго алгоритма кодирования, без фактического кодирования и декодирования части аудиосигнала с использованием второго алгоритма кодирования. Кроме того, устройство 10 содержит контроллер 16 для выбора первого алгоритма кодирования или второго алгоритма 25 кодирования на основе сравнения между первым показателем качества и вторым показателем качества. Контроллер может содержать выход 18, указывающий выбранный алгоритм кодирования.

В нижеприведенном подробном описании, первый модуль оценки использует фильтрованную версию аудиосигнала, т.е. фильтрованную версию части аудиосигнала 30 при оценке первого показателя качества, если фильтр 2, выполненный с возможностью уменьшать амплитуду гармоник, предоставляется и не деактивирован, даже если не явно указано.

В варианте осуществления, первая характеристика, ассоциированная с первым алгоритмом кодирования, лучше подходит для музыкальных и шумоподобных сигналов, 35 а вторая характеристика кодирования, ассоциированная со вторым алгоритмом кодирования, лучше подходит для речевых и переходных сигналов. В вариантах осуществления изобретения, первый алгоритм кодирования представляет собой алгоритм кодирования аудио, такой как алгоритм кодирования с преобразованием, например, алгоритм кодирования с MDCT (модифицированным дискретным косинусным 40 преобразованием), такой как алгоритм кодирования на основе ТСХ (возбуждения по кодированию с преобразованием). Другие алгоритмы кодирования с преобразованием могут быть основаны на FFT-преобразовании либо на любом другом преобразовании или гребенке фильтров. В вариантах осуществления изобретения, второй алгоритм кодирования представляет собой алгоритм кодирования речи, такой как алгоритм 45 кодирования на основе CELP (линейного прогнозирования с возбуждением по коду), к примеру, алгоритм кодирования на основе ACELP (линейного прогнозирования с возбуждением по алгебраическому коду).

В вариантах осуществления, показатель качества представляет показатель

перцепционного качества. Одно значение, которое является оценкой субъективного качества первого алгоритма кодирования, и одно значение, которое является оценкой субъективного качества второго алгоритма кодирования, могут вычисляться. Алгоритм кодирования, который обеспечивает наилучшее оцененное субъективное качество, может выбираться только на основе сравнения этих двух значений. Это отличается от того, что выполняется в AMR-WB+-стандарте, в котором вычисляются множество признаков, представляющих различные характеристики сигнала, а затем применяется классификатор для того, чтобы определять, какой алгоритм следует выбирать.

В вариантах осуществления, соответствующий показатель качества оценивается на основе части взвешенного аудиосигнала, т.е. взвешенной версии аудиосигнала. В вариантах осуществления, взвешенный аудиосигнал может задаваться как аудиосигнал, фильтрованный посредством функции взвешивания, при этом функция взвешивания представляет собой взвешенный LPC-фильтр $A(z/g)$, где $A(z)$ является LPC-фильтром, а g является весовым коэффициентом между 0 и 1, к примеру, 0,68. Оказывается, что хорошие показатели перцепционного качества могут получаться таким способом. Следует отметить, что LPC-фильтр $A(z)$ и взвешенный LPC-фильтр $A(z/g)$ определяются в каскаде предварительной обработки, и что они также используются в обоих алгоритмах кодирования. В других вариантах осуществления, функция взвешивания может представлять собой линейный фильтр, FIR-фильтр или линейный прогнозный фильтр.

В вариантах осуществления, показатель качества представляет собой сегментальное SNR (отношение "сигнал-шум") в области взвешенных сигналов. Оказывается, что сегментальное SNR в области взвешенных сигналов представляет хороший показатель перцепционного качества и в силу этого может использоваться в качестве показателя качества преимущественным способом. Оно также представляет собой показатель качества, используемый в обоих алгоритмах ACELP- и TCX-кодирования для того, чтобы оценивать параметры кодирования.

Другой показатель качества может представлять собой SNR в области взвешенных сигналов. Другие показатели качества могут представлять собой сегментальное SNR, SNR соответствующей части аудиосигнала в области невзвешенных сигналов, т.е. нефильТРованное посредством (взвешенных) LPC-коэффициентов.

Обычно, SNR сравнивает исходные и обработанные аудиосигналы (к примеру, речевые сигналы) последовательно выборочно. Его цель состоит в том, чтобы измерять искажение кодеров на основе формы сигналов, которые воспроизводят форму входного сигнала. SNR может вычисляться так, как показано на фиг. 5а, где $x(i)$ и $y(i)$ являются исходными и обработанными выборками, индексированными посредством i , и N является общим числом выборок. Сегментальное SNR, вместо обработки для всего сигнала, вычисляет среднее SNR-значений коротких сегментов, к примеру, в 1-10 мс, к примеру, в 5 мс. SNR может вычисляться так, как показано на фиг. 5b, где N и M являются длиной сегмента и числом сегментов, соответственно.

В вариантах осуществления изобретения, часть аудиосигнала представляет кадр аудиосигнала, который получается посредством кодирования с взвешиванием аудиосигнала, и выбор надлежащего алгоритма кодирования выполняется для множества последовательных кадров, полученных посредством кодирования со взвешиванием аудиосигнала. В нижеприведенном подробном описании, в связи с аудиосигналом, термины "часть" и "кадр" используются взаимозаменяемо. В вариантах осуществления, каждый кадр разделен на субкадры, и сегментальное SNR оценивается для каждого кадра посредством вычисления SNR для каждого субкадра,

преобразованного в дБ, и вычисления среднего SNR субкадров в дБ.

Таким образом, в вариантах осуществления, оценивается не (сегментальное) SNR между входным аудиосигналом и декодированным аудиосигналом, а оценивается (сегментальное) SNR между взвешенным входным аудиосигналом и взвешенным декодированным аудиосигналом. Что касается этого (сегментального) SNR, можно обратиться к главе 5.2.3 AMR-WB+-стандарта (Международного стандарта 3GPP TS 26.290 V6.1.0 2004-12).

В вариантах осуществления изобретения, соответствующий показатель качества оценивается на основе энергии части взвешенного аудиосигнала и на основе оцененного искажения, введенного при кодировании части сигнала посредством соответствующего алгоритма, при этом первый и второй модули оценки выполнены с возможностью определять оцененные искажения в зависимости от энергии взвешенного аудиосигнала.

В вариантах осуществления изобретения, определяется оцененное искажение квантователя, введенное посредством квантователя, используемого в первом алгоритме кодирования при квантовании части аудиосигнала, и первый показатель качества определяется на основе энергии части взвешенного аудиосигнала и оцененного искажения квантователя. В таких вариантах осуществления, глобальное усиление для части аудиосигнала может оцениваться таким образом, что часть аудиосигнала должна формировать данную целевую скорость передачи битов при кодировании с помощью квантователя и энтропийного кодера, используемых в первом алгоритме кодирования, при этом оцененное искажение квантователя определяется на основе оцененного глобального усиления. В таких вариантах осуществления, оцененное искажение квантователя может определяться на основе мощности оцененного усиления. Когда квантователь, используемый в первом алгоритме кодирования, представляет собой равномерный скалярный квантователь, первый модуль оценки может быть выполнен с возможностью определять оцененное искажение квантователя с использованием формулы $D=G*G/12$, где D является оцененным искажением квантователя, и G является оцененным глобальным усилением. В случае если первый алгоритм кодирования использует другой квантователь, искажение квантователя может определяться из глобального усиления другим способом.

Авторы изобретения выяснили, что показатель качества, такой как сегментальное SNR, которое получается при кодировании и декодировании части аудиосигнала с использованием первого алгоритма кодирования, такого как ТСХ-алгоритм, может оцениваться надлежащим образом посредством использования вышеуказанных признаков в любой их комбинации.

В вариантах осуществления изобретения, первый показатель качества представляет собой сегментальное SNR, и сегментальное SNR оценивается посредством вычисления оцененного SNR, ассоциированного с каждой из множества подчастей части аудиосигнала, на основе энергии соответствующей подчасти взвешенного аудиосигнала и оцененного искажения квантователя и посредством вычисления среднего SNR, ассоциированных с подчастями части взвешенного аудиосигнала, чтобы получать оцененное сегментальное SNR для части взвешенного аудиосигнала.

В вариантах осуществления изобретения, определяется оцененное искажение адаптивной таблицы кодирования, введенное посредством адаптивной таблицы кодирования, используемой во втором алгоритме кодирования при использовании адаптивной таблицы кодирования для того, чтобы кодировать часть аудиосигнала, и второй показатель качества оценивается на основе энергии части взвешенного аудиосигнала и оцененного искажения адаптивной таблицы кодирования.

В таких вариантах осуществления, для каждой из множества подчастей части аудиосигнала, адаптивная таблица кодирования может быть аппроксимирована на основе версии подчасти взвешенного аудиосигнала, сдвинутой в прошлое посредством запаздывания основного тона, определенного в каскаде предварительной обработки, усиление адаптивной таблицы кодирования может оцениваться таким образом, что ошибка между подчастью части взвешенного аудиосигнала и аппроксимированной адаптивной таблицей кодирования минимизируется, и оцененное искажение адаптивной таблицы кодирования может определяться на основе энергии ошибки между подчастью части взвешенного аудиосигнала и аппроксимированной адаптивной таблицей кодирования, масштабированной посредством усиления адаптивной таблицы кодирования.

В вариантах осуществления изобретения, оцененное искажение адаптивной таблицы кодирования, определенное для каждой подчасти части аудиосигнала, может уменьшаться на постоянный множитель, чтобы учитывать уменьшение искажения, которое достигается посредством инновационной таблицы кодирования во втором алгоритме кодирования.

В вариантах осуществления изобретения, второй показатель качества представляет собой сегментальное SNR, и сегментальное SNR оценивается посредством вычисления оцененного SNR, ассоциированного с каждой подчастью, на основе энергии соответствующей подчасти взвешенного аудиосигнала и оцененного искажения адаптивной таблицы кодирования и посредством вычисления среднего SNR, ассоциированных с подчастями, чтобы получать оцененное сегментальное SNR.

В вариантах осуществления изобретения, адаптивная таблица кодирования аппроксимирована на основе версии части взвешенного аудиосигнала, сдвинутой в прошлое посредством запаздывания основного тона, определенного в каскаде предварительной обработки, усиление адаптивной таблицы кодирования оценивается таким образом, что ошибка между частью взвешенного аудиосигнала и аппроксимированной адаптивной таблицей кодирования минимизируется, и оцененное искажение адаптивной таблицы кодирования определяется на основе энергии между частью взвешенного аудиосигнала и аппроксимированной адаптивной таблицей кодирования, масштабированной посредством усиления адаптивной таблицы кодирования. Таким образом, оцененное искажение адаптивной таблицы кодирования может определяться с низкой сложностью.

Авторы изобретения выяснили, что показатель качества, такой как сегментальное SNR, которое получается при кодировании и декодировании части аудиосигнала с использованием второго алгоритма кодирования, такого как ACELP-алгоритм, может оцениваться надлежащим образом посредством использования вышеуказанных признаков в любой их комбинации.

В вариантах осуществления изобретения, гистерезисный механизм используется при сравнении оцененных показателей качества. Он позволяет обеспечивать большую стабильность решения в отношении того, какой алгоритм должен использоваться. Гистерезисный механизм может зависеть от оцененных показателей качества (таких как разность между ними) и других параметров, таких как статистика в отношении предыдущих решений, число временно стационарных кадров, переходные части в кадрах. Что касается таких гистерезисных механизмов, можно обратиться, например, к WO 2012/110448 A1.

В вариантах осуществления изобретения, кодер для кодирования аудиосигнала содержит устройство 10, каскад для выполнения первого алгоритма кодирования и

каскад для выполнения второго алгоритма кодирования, при этом кодер выполнен с возможностью кодировать часть аудиосигнала с использованием первого алгоритма кодирования или второго алгоритма кодирования в зависимости от выбора посредством контроллера 16. В вариантах осуществления изобретения, система для кодирования и декодирования содержит кодер и декодер, выполненные с возможностью принимать кодированную версию части аудиосигнала и индикатор относительно алгоритма, используемого для того, чтобы кодировать часть аудиосигнала и декодировать кодированную версию части аудиосигнала с использованием указываемого алгоритма.

Такой алгоритм выбора режима с разомкнутым контуром, как показано на фиг. 1 и описано выше (за исключением фильтра 2), описывается в предыдущей заявке РСТ/EP2014/051557. Этот алгоритм используется для того, чтобы проводить выбор между двумя режимами, такими как ACELP и TCX, на покадровой основе. Выбор может быть основан на оценке сегментального SNR как ACELP, так и TCX. Режим с наибольшим оцененным сегментированным SNR выбирается. Необязательно, гистерезисный механизм может использоваться для того, чтобы предоставлять более надежный выбор. Сегментальное SNR ACELP может оцениваться с использованием аппроксимации искажения адаптивной таблицы кодирования и аппроксимации искажения инновационной таблицы кодирования. Адаптивная таблица кодирования может быть аппроксимирована в области взвешенных сигналов с использованием запаздывания основного тона, оцененного посредством алгоритма анализа основного тона. Искражение может вычисляться в области взвешенных сигналов при условии оптимального усиления. Искражение затем может уменьшаться посредством постоянного множителя, аппроксимирующего искажение инновационной таблицы кодирования. Сегментальное SNR TCX может оцениваться с использованием упрощенной версии реального TCX-кодера. Входной сигнал может сначала преобразовываться с помощью MDCT и затем формироваться с использованием взвешенного LPC-фильтра. В завершение, искажение может оцениваться во взвешенной MDCT-области, с использованием глобального усиления и модуля оценки глобального усиления.

Оказывается, что этот алгоритм выбора режима с разомкнутым контуром, как описано в предыдущей заявке, предоставляет ожидаемое решение большую часть времени, выбирая ACELP для речевых и переходных сигналов и TCX для музыкальных и шумоподобных сигналов. Тем не менее, авторы изобретения выяснили, что может возникать такая ситуация, что ACELP иногда выбирается для некоторых гармонических музыкальных сигналов. Для таких сигналов, адаптивная таблица кодирования, в общем, имеет высокое усиление для прогнозирования вследствие высокой прогнозируемости гармонических сигналов, формируя низкое искажение и затем более высокое сегментальное SNR по сравнению с TCX. Тем не менее, TCX звучит лучше для большинства гармонических музыкальных сигналов, так что TCX должен быть предпочтительным в этих случаях.

Таким образом, настоящее изобретение предлагает выполнять оценку SNR или сегментального SNR в качестве первого показателя качества с использованием версии входного сигнала, которая фильтруется для того, чтобы уменьшать ее гармоники. Таким образом, может получаться улучшенный выбор режима для гармонических музыкальных сигналов.

Обычно, может использоваться любой подходящий фильтр для уменьшения гармоник. В вариантах осуществления изобретения, фильтр представляет собой фильтр долговременного прогнозирования. Один простой пример фильтра долговременного прогнозирования следующий:

$$F(z)=1-g*z^{-T},$$

где параметры фильтрации являются усилением g и запаздыванием T основного тона, которые определяются из аудиосигнала.

Варианты осуществления изобретения основаны на фильтре долговременного прогнозирования, который применяется к аудиосигналу перед MDCT-анализом при оценке сегментального TCX SNR. Фильтр долговременного прогнозирования уменьшает амплитуду гармоник во входном сигнале перед MDCT-анализом. Последствие состоит в том, что искажение во взвешенной MDCT-области уменьшается, оцененное сегментальное SNR TCX увеличивается, и в завершение, TCX выбирается чаще для гармонических музыкальных сигналов.

В вариантах осуществления изобретения, передаточная функция фильтра долговременного прогнозирования содержит целочисленную часть запаздывания основного тона и многоотводного фильтра в зависимости от дробной части запаздывания основного тона. Это обеспечивает эффективную реализацию, поскольку целочисленная часть используется только в нормальной концепции ($z^{-T_{\text{int}}}$) частоты дискретизации. Одновременно, может достигаться высокая точность вследствие использования дробной части в многоотводном фильтре. Посредством рассмотрения дробной части в многоотводном фильтре, может достигаться удаление энергии гармоник, тогда как удаление энергии частей около гармоник не допускается.

В вариантах осуществления изобретения, фильтр долговременного прогнозирования описывается следующим образом:

$$P(z) = 1 - \beta g B(z, T_{fr}) z^{-T_{\text{int}}},$$

где T_{int} и T_{fr} являются целочисленной и дробной частью запаздывания основного тона, g усиление, β является весовым коэффициентом, и $B(z, T_{fr})$ является FIR-фильтром нижних частот, коэффициенты которого зависят от дробной части запаздывания основного тона. Ниже изложена более подробная информация касательно вариантов осуществления такого фильтра долговременного прогнозирования.

Запаздывание и усиление основного тона могут оцениваться на покадровой основе.

Прогнозный фильтр может деактивироваться (усиление = 0) на основе комбинации одного или более показателей гармонического характера (например, нормализованной корреляции или усиления для прогнозирования) и/или одного или более показателей временной структуры (например, показателя временной равномерности или изменения энергии).

Фильтр может применяться к входному аудиосигналу на покадровой основе. Если параметры фильтрации изменяются между кадрами, разрыв может вводиться на границе между двумя кадрами. В вариантах осуществления, устройство дополнительно содержит модуль для удаления разрывов в аудиосигнале, вызываемых посредством фильтра.

Чтобы удалять возможные разрывы, может использоваться любая технология, к примеру, технологии, сравнимые с технологиями, описанными в US5012517, EP0732687A2, US5999899A или US7353168B2. Ниже описывается другая технология для удаления возможных разрывов.

Перед подробным описанием варианта осуществления первого модуля 12 оценки и второго модуля 14 оценки со ссылкой на фиг. 3, описывается вариант осуществления кодера 20 со ссылкой на фиг. 2.

Кодер 20 содержит первый модуль 12 оценки, второй модуль 14 оценки, контроллер 16, препроцессор 22, переключатель 24, первый каскад 26 кодера, выполненный с возможностью осуществлять TCX-алгоритм, второй каскад 28 кодера, выполненный

с возможностью осуществлять ACELP-алгоритм, и интерфейс 30 вывода. Препроцессор 22 может быть частью общего USAC-кодера и может быть выполнен с возможностью выводить LPC-коэффициенты, взвешенные LPC-коэффициенты, взвешенный аудиосигнал и набор запаздываний основного тона. Следует отметить, что все эти параметры

5 используются в обоих алгоритмах кодирования, т.е. в TCX-алгоритме и в ACELP-алгоритме. Таким образом, такие параметры не должны дополнительно вычисляться для решения по выбору режима с разомкнутым контуром. Преимущество использования уже вычисленных параметров в решении по выбору режима с разомкнутым контуром состоит в сокращении сложности.

10 Как показано на фиг. 2, устройство содержит фильтр 2 уменьшения гармоник. Устройство дополнительно содержит необязательный модуль 4 деактивации для деактивации фильтра 2 уменьшения гармоник на основе комбинации одного или более показателей гармонического характера (например, нормализованной корреляции или усиления для прогнозирования) и/или одного или более показателей временной

15 структуры (например, показателя временной равномерности или изменения энергии). Устройство содержит необязательный модуль 6 удаления разрывов для удаления разрывов из фильтрованной версии аудиосигнала. Помимо этого, устройство необязательно содержит модуль 8 для оценки параметров фильтрации фильтра 2 уменьшения гармоник. На фиг. 2, эти компоненты (2, 4, 6 и 8) показаны как часть

20 первого модуля 12 оценки. Само собой разумеется, что эти компоненты могут реализовываться внешне или отдельно от первого модуля оценки и могут быть выполнены с возможностью предоставлять фильтрованную версию аудиосигнала в первый модуль оценки.

Входной аудиосигнал 40 предоставляется во входной линии. Входной аудиосигнал

25 40 применяется к первому модулю 12 оценки, препроцессору 22 и обоим каскадам 26, 28 кодера. В первом модуле 12 оценки, входной аудиосигнал 40 применяется к фильтру 2, и фильтрованная версия входного аудиосигнала используется при оценке первого показателя качества. В случае если фильтр деактивируется посредством модуля 4 деактивации, входной аудиосигнал 40 используется при оценке первого показателя

30 качества, а не фильтрованной версии входного аудиосигнала. Препроцессор 22 обрабатывает входной аудиосигнал традиционным способом, чтобы извлекать LPC-коэффициенты и взвешенные LPC-коэффициенты 42 и фильтровать аудиосигнал 40 со взвешенными LPC-коэффициентами 42, чтобы получать взвешенный аудиосигнал 44. Препроцессор 22 выводит взвешенные LPC-коэффициенты 42, взвешенный аудиосигнал

35 44 и набор запаздываний 48 основного тона. Специалисты в данной области техники должны понимать, что взвешенные LPC-коэффициенты 42 и взвешенный аудиосигнал 44 могут быть сегментированы на кадры или субкадры. Сегментация может получаться посредством кодирования с взвешиванием аудиосигнала надлежащим образом.

В альтернативных вариантах осуществления, может предоставляться препроцессор,

40 который выполнен с возможностью формировать взвешенные LPC-коэффициенты и взвешенный аудиосигнал на основе фильтрованной версии аудиосигнала. Взвешенные LPC-коэффициенты и взвешенный аудиосигнал, которые основаны на фильтрованной версии аудиосигнала, затем применяются к первому модулю оценки, чтобы оценивать первый показатель качества, вместо взвешенных LPC-коэффициентов 42 и взвешенного

45 аудиосигнала 44.

В вариантах осуществления изобретения, могут использоваться квантованные LPC-коэффициенты или квантованные взвешенные LPC-коэффициенты. Таким образом, следует понимать, что термин "LPC-коэффициенты" также имеет намерение охватывать

"квантованные LPC-коэффициенты", а термин "взвешенные LPC-коэффициенты" также имеет намерение охватывать "взвешенные квантованные LPC-коэффициенты". В этом отношении, необходимо отметить, что TCX-алгоритм USAC использует квантованные взвешенные LPC-коэффициенты, чтобы формировать MCDT-спектр.

5 Первый модуль 12 оценки принимает аудиосигнал 40, взвешенные LPC-коэффициенты 42 и взвешенный аудиосигнал 44, оценивает первый показатель 46 качества на их основе и выводит первый показатель качества в контроллер 16. Второй модуль 16 оценки принимает взвешенный аудиосигнал 44 и набор запаздываний 48 основного тона, оценивает второй показатель 50 качества на их основе и выводит второй показатель
10 50 качества в контроллер 16. Как известно специалистам в данной области техники, взвешенные LPC-коэффициенты 42, взвешенный аудиосигнал 44 и набор запаздываний 48 основного тона уже вычислены в предыдущем модуле (т.е. в препроцессоре 22) и в силу этого доступны без вычислительных затрат.

Контроллер принимает решение, чтобы выбирать либо TCX-алгоритм, либо ACELP-
15 алгоритм, на основе сравнения принимаемых показателей качества. Как указано выше, контроллер может использовать гистерезисный механизм при определении того, какой алгоритм должен использоваться. Выбор первого каскада 26 кодера или второго каскада 28 кодера схематично показан на фиг. 2 посредством переключателя 24, который управляется посредством управляющего сигнала 52, выводимого посредством
20 контроллера 16. Управляющий сигнал 52 указывает то, должен использоваться первый каскад 26 кодера или второй каскад 28 кодера. На основе управляющего сигнала 52, требуемые сигналы, схематично указываемые посредством стрелки 54 на фиг. 2 и включающие в себя, по меньшей мере, LPC-коэффициенты, взвешенные LPC-
коэффициенты, аудиосигнал, взвешенный аудиосигнал, набор запаздываний основного
25 тона, применяются либо к первому каскаду 26 кодера, либо к второму каскаду 28 кодера. Выбранный каскад кодера применяет ассоциированный алгоритм кодирования и выводит кодированное представление 56 или 58 в интерфейс 30 вывода. Интерфейс 30 вывода может быть выполнен с возможностью выводить кодированный аудиосигнал 60, который может содержать, в числе других данных, кодированное представление 56
30 или 58, LPC-коэффициенты или взвешенные LPC-коэффициенты, параметры для выбранного алгоритма кодирования и информацию относительно выбранного алгоритма кодирования.

Далее описываются конкретные варианты осуществления для оценки первого и второго показателей качества, при этом первый и второй показатели качества являются
35 сегментальными SNR в области взвешенных сигналов, со ссылкой на фиг. 3. Фиг. 3 показывает первый модуль 12 оценки и второй модуль 14 оценки и их функциональности в форме блок-схем последовательности операций способа, пошагово показывающих соответствующую оценку.

Оценка сегментального TCX SNR

40 Первый модуль оценки (TCX) принимает аудиосигнал 40 (входной сигнал), взвешенные LPC-коэффициенты 42 и взвешенный аудиосигнал 44 в качестве вводов. Фильтрованная версия аудиосигнала 40 формируется, этап 98. В фильтрованной версии аудиосигнала 40 гармоники уменьшаются или подавляются.

Аудиосигнал 40 может анализироваться для того, чтобы определять один или более
45 показателей гармонического характера (например, нормализованную корреляцию или усиление для прогнозирования) и/или один или более показателей временной структуры (например, показатель временной равномерности или изменение энергии). На основе одного из этих показателей или комбинации этих показателей, может деактивироваться

фильтр 2 и в силу этого фильтрация 98. Если фильтрация 98 деактивируется, оценка первого показателя качества выполняется с использованием аудиосигнала 40, а не его фильтрованной версии.

В вариантах осуществления изобретения, этап удаления разрывов (не показан на фиг. 3) может выполняться после фильтрации 98, чтобы удалять разрывы в аудиосигнале, которые могут получаться в результате фильтрации 98.

На этапе 100, фильтрованная версия аудиосигнала 40 кодируется с взвешиванием. Кодирование со взвешиванием может осуществляться с синусоидальным окном с низким перекрытием в 10 мс. Когда предыдущий кадр представляет собой ACELP, размер блока может увеличиваться на 5 мс, левая сторона окна может быть прямоугольной, и кодированная со взвешиванием нулевая импульсная характеристика синтезирующего ACELP-фильтра может быть удалена из кодированного со взвешиванием входного сигнала. Это является аналогичным тому, что выполняется в ТСХ-алгоритме. Кадр фильтрованной версии аудиосигнала 40, который представляет часть аудиосигнала, выводится из этапа 100.

На этапе 102, кодированный со взвешиванием аудиосигнал, т.е. результирующий кадр, преобразуется с помощью MDCT (модифицированного дискретного косинусного преобразования). На этапе 104, формирование спектра выполняется посредством формирования MDCT-спектра с взвешенными LPC-коэффициентами.

На этапе 106, глобальное усиление G оценивается таким образом, что взвешенный спектр, квантованный с усилением G , должен формировать данный целевой R при кодировании с помощью энтропийного кодера, например, арифметического кодера. Термин "глобальное усиление" используется, поскольку одно усиление определяется для целого кадра.

Далее поясняется пример реализации оценки глобального усиления. Следует отметить, что эта оценка глобального усиления является подходящей для вариантов осуществления, в которых алгоритм ТСХ-кодирования использует скалярный квантователь с арифметическим кодером. Такой скалярный квантователь с арифметическим кодером предполагается в MPEG USAC-стандарте.

Инициализация

Во-первых, переменные, используемые при оценке усиления, инициализируются следующим образом:

set $en[i]=9,0+10,0*\log_{10}(c[4*i+0]+c[4*i+1]+c[4*i+2]+c[4*i+3])$,

где $0 \leq i < L/4$, $c[]$ является вектором коэффициентов для квантования, и L является длиной $c[]$.

2. set $fac=128$, $offset=fac$ and $target=any\ value\ (e.g.\ 1000)$

Итерация

Затем следующий блок операций выполняется NITER раз (например, здесь, NITER=10).

1. $fac=fac/2$

2. $offset=offset-fac$

3. $ener=0$

4. for every i where $0 \leq i < L/4$ do the following:

if $en[i]-offset > 3.0$, then $ener=ener+en[i]-offset$

5. if $ener > target$, then $offset=offset+fac$

Результат итерации представляет собой значение смещения. После итерации глобальное усиление оценивается как $G=10^{(offset/20)}$.

Конкретный способ, которым оценивается глобальное усиление, может варьироваться

в зависимости от используемых квантователя и энтропийного кодера. В MPEG USAC-стандарте предполагается скалярный квантователь с арифметическим кодером. Другие ТСХ-подходы могут использовать другой квантователь, и специалисты в данной области техники должны понимать, как оценивать глобальное усиление для таких других квантователей. Например, AMR-WB+-стандарт предполагает то, что используется решетчатый RE8-квантователь. Для такого квантователя, оценка глобального усиления может оцениваться так, как описано в главе 5.3.5.7 на странице 34 3GPP TS 26.290 V6.1.0 2004-12, в которой предполагается фиксированная целевая скорость передачи битов.

После оценки глобального усиления на этапе 106, оценка искажения осуществляется на этапе 108. Более конкретно, искажение квантователя аппроксимировано на основе оцененного глобального усиления. В настоящем варианте осуществления предполагается, что используется равномерный скалярный квантователь. Таким образом, искажение квантователя определяется с помощью простой формулы $D=G*G/12$, в которой D представляет определенное искажение квантователя, а G представляет оцененное глобальное усиление. Это соответствует высокоскоростной аппроксимации искажения равномерного скалярного квантователя.

На основе определенного искажения квантователя вычисление сегментального SNR выполняется на этапе 110. SNR в каждом субкадре кадра вычисляется как отношение энергии взвешенного аудиосигнала и искажения D , которое предполагается постоянным в субкадрах. Например, кадр разбивается на четыре последовательных субкадра (см. фиг. 4). Сегментальное SNR затем является средним SNR четырех субкадров и может указываться в дБ.

Этот подход обеспечивает оценку первого сегментального SNR, которая получается при фактическом кодировании и декодировании настоящего кадра с использованием ТСХ-алгоритма, однако без необходимости фактически кодировать и декодировать аудиосигнал и в силу этого со значительно меньшей сложностью и уменьшенным временем вычислений.

Оценка сегментального ACELP SNR

Второй модуль 14 оценки принимает взвешенный аудиосигнал 44 и набор запаздываний 48 основного тона, который уже вычислен в препроцессоре 22.

Как показано на этапе 112, в каждом субкадре, адаптивная таблица кодирования аппроксимирована посредством простого использования взвешенного аудиосигнала и запаздывания T основного тона. Адаптивная таблица кодирования аппроксимирована следующим образом:

$$xw(n-T), n=0, \dots, N,$$

где xw является взвешенным аудиосигналом, T является запаздыванием основного тона соответствующего субкадра, и N является длиной субкадра. Соответственно, адаптивная таблица кодирования аппроксимирована посредством использования версии субкадра, сдвинутой в прошлое на T . Таким образом, в вариантах осуществления изобретения, адаптивная таблица кодирования аппроксимирована очень простым способом.

На этапе 114, определяется усиление адаптивной таблицы кодирования для каждого субкадра. Более конкретно, в каждом субкадре, усиление G таблицы кодирования оценивается таким образом, что оно минимизирует ошибку между взвешенным аудиосигналом и аппроксимированной адаптивной таблицей кодирования. Это может выполняться посредством простого сравнения разностей между обоими сигналами для каждой выборки и нахождения усиления таким образом, что сумма этих разностей является минимальной.

На этапе 116, определяется искажение адаптивной таблицы кодирования для каждого субкадра. В каждом субкадре, искажение D, введенное посредством адаптивной таблицы кодирования, является просто энергией ошибки между взвешенным аудиосигналом и аппроксимированной адаптивной таблицей кодирования, масштабированной

5 посредством усиления G.

Искажения, определенные на этапе 116, могут регулироваться на необязательном этапе 118, с тем, чтобы учитывать инновационную таблицу кодирования. Искажение инновационной таблицы кодирования, используемой в ACELP-алгоритмах, может быть просто оценено в качестве постоянного значения. В описанном варианте осуществления

10 изобретения, просто предполагается, что инновационная таблица кодирования уменьшает искажение D посредством постоянного множителя. Таким образом, искажения, полученные на этапе 116 для каждого субкадра, могут умножаться на этапе 118 посредством постоянного множителя, к примеру, постоянный множитель в порядке от 0 до 1, к примеру, 0,055.

На этапе 120, осуществляется вычисление сегментального SNR. В каждом субкадре SNR вычисляется как отношение энергии взвешенного аудиосигнала и искажения D. Сегментальное SNR является затем средним значением SNR четырех субкадров и может указываться в дБ.

Этот подход обеспечивает оценку второго SNR, которая получается при фактическом

20 кодировании и декодировании настоящего кадра с использованием ACELP-алгоритма, однако без необходимости фактически кодировать и декодировать аудиосигнал и, поэтому, со строго меньшей сложностью и уменьшенным временем вычислений.

Первый и второй модули 12 и 14 оценки выводят оцененные сегментальные SNR 46, 50 в контроллер 16, и контроллер 16 принимает решение в отношении того, какой

25 алгоритм должен использоваться для ассоциированной части аудиосигнала, на основе оцененных сегментальных SNR 46, 50. Контроллер необязательно может использовать гистерезисный механизм, чтобы принимать более стабильное решение. Например, гистерезисный механизм, идентичный гистерезисному механизму в решении с замкнутым контуром, может использоваться с немного отличающимися параметрами настройки.

Такой гистерезисный механизм может вычислять значение "dsnr", которое может зависеть от оцененных сегментальных SNR (таких как разность между ними) и других параметров, таких как статистика в отношении предыдущих решений, число временно стационарных кадров и переходные части в кадрах.

Без гистерезисного механизма, контроллер может выбирать алгоритм кодирования,

35 имеющий более высокое оцененное SNR, т.е. ACELP выбирается, если второе оцененное SNR выше первого оцененного SNR, а TCX выбирается, если первое оцененное SNR выше второго оцененного SNR. При использовании гистерезисного механизма, контроллер может выбирать алгоритм кодирования согласно следующему правилу принятия решения, где acelp_snr является вторым оцененным SNR, и tcx_snr является

40 первым оцененным SNR:

if acelp_snr+dsnr>tcx_snr then select ACELP, otherwise select TCX.

Определение параметров фильтра для уменьшения амплитуды гармоник

Далее описывается вариант осуществления для определения параметров фильтра для уменьшения амплитуды гармоник. Параметры фильтрации могут оцениваться на

45 стороне кодера, к примеру, в модуле 8.

Оценка основного тона

Одно запаздывание основного тона (целочисленная часть + дробная часть) в расчете на кадр оценивается (размер кадра, например, 20 мс). Это выполняется на трех этапах

для того, чтобы уменьшать сложность и повышать точность оценки.

а) Первая оценка целочисленной части запаздывания основного тона

Алгоритм анализа основного тона, который формирует гладкий контур изменения основного тона, используется (например, анализ основного тона с разомкнутым контуром, описанный в Rec. ITU-T G.718, раздел 6.6). Этот анализ, в общем, выполняется на основе субкадров (размер субкадра, например, 10 мс) и формирует одну оценку запаздывания основного тона в расчете на субкадр. Следует отметить, что эти оценки запаздывания основного тона не имеют дробной части и, в общем, оцениваются для сигнала после понижающей дискретизации (с частотой дискретизации, например, 6400 Гц). Используемый сигнал может представлять собой любой аудиосигнал, например, взвешенный LPC-аудиосигнал, как описано в Rec. ITU-T G.718, раздел 6.5.

б) Детализация целочисленной части T_{int} запаздывания основного тона

Конечная целочисленная часть запаздывания основного тона оценивается для аудиосигнала $x[n]$, работающего на частоте дискретизации базового кодера, которая, в общем, выше частоты дискретизации сигнала после понижающей дискретизации, используемого в а) (например, 12,8 кГц, 16 кГц, 32 кГц, ...). Сигнал $x[n]$ может представлять собой любой аудиосигнал, например, взвешенный LPC-аудиосигнал.

Целочисленная часть T_{int} запаздывания основного тона в таком случае является запаздыванием, которое максимизирует автокорреляционную функцию

$$C(d) = \sum_{n=0}^N x[n]x[n-d],$$

где d вокруг запаздывания T основного тона оценено в а).

$$T - \delta_1 \leq d \leq T + \delta_2$$

с) Оценка дробной части T_{fr} запаздывания основного тона

Дробная часть T_{fr} находится посредством интерполяции автокорреляционной функции $C(d)$, вычисленной на этапе б), и выбора дробного запаздывания основного тона, которое максимизирует интерполированную автокорреляционную функцию.

Интерполяция может выполняться с использованием FIR-фильтра нижних частот, как описано в, например, Rec. ITU-T G.718, раздел 6.6.7.

Оценка и квантование усиления

Усиление, в общем, оценивается для входного аудиосигнала на частоте дискретизации базового кодера, но он также может представлять собой любой аудиосигнал, к примеру, взвешенный LPC-аудиосигнал. Этот сигнал помечен $y[n]$ и может быть идентичным или отличающимся от $x[n]$.

Прогнозирование $yP[n]$ $y[n]$ сначала находится посредством фильтрации $y[n]$ со следующим фильтром:

$$P(z) = B(z, T_{fr})z^{-T_{int}},$$

где T_{int} является целочисленной частью запаздывания основного тона (оцененной в б)), и $B(z, T_{fr})$ является FIR-фильтром нижних частот, коэффициенты которого зависят от дробной части запаздывания основного тона T_{fr} (оцененным в с)).

Один пример $B(z)$, когда разрешение запаздывания основного тона составляет $\frac{1}{4}$:

$$\begin{aligned}
 T_{fr} = \frac{0}{4} & \quad B(z) = 0.0000z^{-2} + 0.2325z^{-1} + 0.5349z^0 + 0.2325z^1 \\
 T_{fr} = \frac{1}{4} & \quad B(z) = 0.0152z^{-2} + 0.3400z^{-1} + 0.5094z^0 + 0.1353z^1 \\
 T_{fr} = \frac{2}{4} & \quad B(z) = 0.0609z^{-2} + 0.4391z^{-1} + 0.4391z^0 + 0.0609z^1 \\
 T_{fr} = \frac{3}{4} & \quad B(z) = 0.1353z^{-2} + 0.5094z^{-1} + 0.3400z^0 + 0.0152z^1
 \end{aligned}$$

Усиление g затем вычислено следующим образом:

$$g = \frac{\sum_{n=0}^{N-1} y[n]y_p[n]}{\sum_{n=0}^{N-1} y_p[n]y_p[n]}$$

и ограничено между 0 и 1.

В завершение, усиление g квантуется, например, в 2 битах, с использованием, например, равномерного квантования.

β используется для того, чтобы управлять интенсивностью фильтра. β , равный 1, формирует полное действие. β , равный 0, деактивирует фильтр. Таким образом, в вариантах осуществления изобретения, фильтр может деактивироваться посредством задания β равным значению 0. В вариантах осуществления изобретения, если фильтр активируется, β может задаваться равным значению между 0,5 и 0,75. В вариантах осуществления изобретения, если фильтр активируется, β может задаваться равным значению 0,625. Пример $B(z, T_{fr})$ приведен выше. Порядок и коэффициенты $B(z, T_{fr})$ также могут зависеть от скорости передачи битов и выходной частоты дискретизации. Различная частотная характеристика может быть рассчитана и настроена для каждой комбинации скорости передачи битов и выходной частоты дискретизации.

Деактивация фильтра

Фильтр может деактивироваться на основе комбинации одного или более показателей гармонического характера и/или одного или более показателей временной структуры. Примеры таких показателей описываются ниже:

i) Показатель гармонического характера, такой как нормализованная корреляция при целочисленном запаздывании основного тона, оцененном на этапе b).

$$norm.corr. = \frac{\sum_{n=0}^N x[n]x[n - T_{int}]}{\sqrt{\sum_{n=0}^N x[n]x[n]} \sqrt{\sum_{n=0}^N x[n - T_{int}]x[n - T_{int}]}}$$

Нормализованная корреляция равна 1, если входной сигнал является идеально прогнозируемым посредством целочисленного запаздывания основного тона, и равна 0, если он вообще не является прогнозируемым. Высокое значение (близко к 1) в таком случае указывает гармонический сигнал. Для более надежного решения, нормализованная корреляция предыдущего кадра также может использоваться в решении, например:

Если $(norm.corr(curr.) * norm.corr(prev.)) > 0,25$, то фильтр не деактивируется

ii) Показатели временной структуры вычисляемые, например, на основе энергетических выборок, также используемые посредством детектора переходных частей для обнаружения переходных частей (например, показатель временной равномерности, изменение энергии), например, если (показатель временной равномерности $> 3,5$, или изменение энергии $> 3,5$), то фильтр деактивируется.

Ниже изложены дополнительные сведения относительно определения одного или более показателей гармонического характера.

Показатель гармонического характера, например, вычислен посредством

нормализованной корреляции аудиосигнала или его предварительной модифицированной версии при или около запаздывания основного тона. Запаздывание основного тона может определяться даже в каскадах, содержащих первый каскад и второй каскад, при этом в первом каскаде, предварительная оценка запаздывания основного тона определяется в области после понижающей дискретизации первой частоты дискретизации, а во втором каскаде, предварительная оценка запаздывания основного тона уточняется на второй частоте дискретизации, превышающей первую частоту дискретизации. Запаздывание основного тона, например, определено с использованием автокорреляции. По меньшей мере, один показатель временной структуры, например, определен во временной области, временно размещенной в зависимости от информации основного тона. Направленный в прошлое во времени конец временной области, например, размещен в зависимости от информации основного тона. Временной направленный в прошлое конец временной области может быть размещен таким образом, что направленный в прошлое во времени конец временной области смещается в направлении в прошлое на временную величину, монотонно увеличивающуюся с увеличением информации основного тона. Направленный в будущее во времени конец временной области может позиционироваться в зависимости от временной структуры аудиосигнала в возможном варианте временной области, идущем от направленного в прошлое во времени конца временной области либо области более высокого влияния на определение показателя временной структуры, до направленного в будущее во времени конца текущего кадра. Амплитуда или отношение между выборками с максимальной и минимальной энергией в возможном варианте временной области может использоваться с этой целью. Например, по меньшей мере, один показатель временной структуры может измерять изменение средней или максимальной энергии аудиосигнала во временной области, и условие деактивации может удовлетворяться, если как, по меньшей мере, один показатель временной структуры меньше предварительно определенного первого порогового значения, так и показатель гармонического характера, для текущего кадра и/или предыдущего кадра, выше второго порогового значения. Условие также удовлетворяется, если показатель гармонического характера, для текущего кадра, выше третьего порогового значения, и показатель гармонического характера, для текущего кадра и/или предыдущего кадра, выше четвертого порогового значения, которое снижается с увеличением запаздывания основного тона.

Далее представлено пошаговое описание конкретного варианта осуществления для определения показателей.

Этап 1. Обнаружение переходных частей и временные показатели

Входной сигнал $S_{HP}(n)$ вводится в детектор переходных частей во временной области. Входной сигнал $S_{HP}(n)$ фильтрован по верхним частотам. Передаточная функция НР-фильтра обнаружения переходных частей задается посредством:

$$H_{TD}(z) = 0.375 - 0.5z^{-1} + 0.125z^{-2} \quad (1)$$

Сигнал, фильтрованный посредством НР-фильтра обнаружения переходных частей,

обозначается как $s_{TD}(n)$. НР-фильтрованный сигнал $s_{TD}(n)$

5 сегментирован на 8 последовательных сегментов идентичной длины. Энергия НР-

фильтрованного сигнала $s_{HP}(n)$ для каждого сегмента вычисляется следующим

образом:

$$10 \quad E_{TD}(i) = \sum_{n=0}^{L_{segment}-1} (s_{TD}(iL_{segment} + n))^2, \quad i = 0, \dots, 7, \quad (2)$$

где $L_{segment} = \frac{L}{8}$ является числом выборок в сегменте в 2,5 миллисекунды на входной частоте дискретизации.

15 Накопленная энергия вычисляется с использованием следующего:

$$E_{Acc} = \max(E_{TD}(i-1), 0.8125E_{Acc}) \quad (3)$$

Атака обнаруживается, если энергия сегмента $E_{TD}(i)$ превышает

20 накопленную энергию на постоянный множитель $attackRatio = 8.5$, и $attackIndex$ задается

равным i :

$$25 \quad E_{TD}(i) > attackRatio \cdot E_{Acc} \quad (4)$$

Если атака не обнаруживается на основе вышеуказанных критериев, но сильное

увеличение энергии обнаруживается в сегменте i , $attackIndex$ задается равным i без

указания присутствия атаки; $attackIndex$ по существу задается как позиция последней атаки в кадре с некоторыми дополнительными ограничениями.

Изменение энергии для каждого сегмента вычисляется следующим образом:

$$35 \quad E_{chng}(i) = \begin{cases} \frac{E_{TD}(i)}{E_{TD}(i-1)}, & E_{TD}(i) > E_{TD}(i-1) \\ \frac{E_{TD}(i-1)}{E_{TD}(i)}, & E_{TD}(i-1) > E_{TD}(i) \end{cases} \quad (5)$$

Показатель временной равномерности вычисляется следующим образом:

$$40 \quad TFM(N_{past}) = \frac{1}{8 + N_{past}} \sum_{i=-N_{past}}^7 E_{chng}(i) \quad (6)$$

Максимальное изменение энергии вычисляется следующим образом:

$$45 \quad MEC(N_{past}, N_{new}) = \max(E_{chng}(-N_{past}), E_{chng}(-N_{past}+1), \dots, E_{chng}(N_{new}-1)) \quad (7)$$

Если индекс $E_{chng}(i)$ ИЛИ $E_{TD}(i)$ является отрицательным, то он указывает значение из предыдущего сегмента с индексацией сегментов относительно текущего кадра.

N_{past} является числом сегментов из предыдущих кадров. Оно равно 0, если

показатель временной равномерности вычисляется для использования в ACELP/TCX-решении. Если показатель временной равномерности вычисляется для TCX LTP-решения, то она равен:

$$N_{past} = 1 + \min \left(8, \left\lceil 8 \frac{pitch}{L} + 0.5 \right\rceil \right) \quad (8)$$

N_{new} является числом сегментов из текущего кадра. Оно равно 8 для непереходных кадров. Для переходных кадров сначала находятся местоположения сегментов с максимальной и минимальной энергией:

$$i_{\max} = \arg \max_{i \in \{N_{past}, \dots, 7\}} E_{TD}(i) \quad (9)$$

$$i_{\min} = \arg \min_{i \in \{N_{past}, \dots, 7\}} E_{TD}(i) \quad (10)$$

Если $E_{TD}(i_{\min}) > 0.375 E_{TD}(i_{\max})$, то N_{new} задается равным $i_{\max} - 3$,

иначе N_{new} задается равным 8.

Этап 2. Переключение длины блока преобразования

Длина перекрытия и длина блока преобразования TCX зависят от наличия переходной части и ее местоположения.

Таблица 1				
Кодирование перекрытия и длины преобразования на основе позиции переходной части				
Индекс атаки	Перекрытие с первым окном следующего кадра	Решение по выбору короткого/длинного преобразования (закодировано в двоичной форме) 0 - длинное, 1 - короткое	Двоичный код для ширины перекрытия	Код перекрытия
Без	ALDO	0	0	00
-2	полное	1	0	10
-1	полное	1	0	10
0	полное	1	0	10
1	полный	1	0	10
2	минимальное	1	10	110
3	половинное	1	11	111
4	половинное	1	11	111
5	минимальное	1	10	110
6	минимальное	0	10	010
7	половинное	0	11	011

Детектор переходных частей, описанный выше, по существу возвращает индекс из последней атаки с таким ограничением, что если имеется несколько переходных частей, то минимальное перекрытие предпочитается по сравнению с половинным перекрытием, которое предпочитается по сравнению с полным перекрытием. Если атака в позиции 2 или 6 не является достаточно сильной, то половинное перекрытие выбирается вместо минимального перекрытия.

Этап 3. Оценка основного тона

Одно запаздывание основного тона (целочисленная часть + дробная часть) в расчете на кадр оценивается (размер кадра, например, 20 мс), как изложено выше на 3 этапах а)-с), для того чтобы уменьшать сложность и повышать точность оценки.

Этап 4. Решающий бит

5 Если входной аудиосигнал не включает в себя содержания гармоник, либо если технология на основе прогнозирования вводит искажения во временной структуре (например, повторение короткой переходной части), то принимается решение, что фильтр деактивируется.

Решение принимается на основе нескольких параметров, таких как нормализованная корреляция при целочисленном запаздывании основного тона и показатели временной структуры.

Нормализованная корреляция при целочисленном запаздывании norm_corr основного тона оценивается так, как изложено выше. Нормализованная корреляция равна 1, если входной сигнал является идеально прогнозируемым посредством целочисленного запаздывания основного тона, и равна 0, если он вообще не является прогнозируемым. 15 Высокое значение (близко к 1) в таком случае указывает гармонический сигнал. Для более надежного решения, помимо нормализованной корреляции для текущего кадра ($\text{norm_corr}(\text{curr})$), в решении также может использоваться нормализованная корреляция предыдущего кадра ($\text{norm_corr}(\text{prev})$), например:

20 $\text{if}(\text{norm_corr}(\text{curr}) * \text{norm_corr}(\text{prev})) > 0,25$
или
 $\text{if} \max(\text{norm_corr}(\text{curr}), \text{norm_corr}(\text{prev})) > 0,5,$
то текущий кадр включает в себя некоторое содержание гармоник.

Показатели временной структуры могут вычисляться посредством детектора 25 переходных частей (например, показатель временной равномерности (уравнение (6)) и максимальное изменение энергии (уравнение 7)), чтобы не допускать активации фильтра для сигнала, содержащего сильную переходную часть или большие временные изменения. Временные признаки вычисляются для сигнала, содержащего текущий кадр

30 (N_{new} сегментов) и предыдущий кадр вплоть до запаздывания основного тона (

N_{past} сегментов). Для ступенчатых переходных частей, которые медленно

35 затухают, все или некоторые признаки вычисляются только до местоположения переходной части ($i_{\text{max}} - 3$), поскольку искажения в негармонической части спектра, введенные посредством ЛТР-фильтрации, должны подавляться посредством маскирования сильной длительной переходной части (например, подвесной тарелки).

40 Последовательности импульсов для низкотональных сигналов могут обнаруживаться в качестве переходной части посредством детектора переходных частей. Для сигналов с низким основным тоном в силу этого игнорируются признаки из детектора переходных частей, и вместо этого предусмотрено дополнительное пороговое значение для нормализованной корреляции, которое зависит от запаздывания основного тона, 45 например:

Если $\text{norm_corr} \leq 1,2 - T_{\text{int}}/L$, то деактивация фильтра

Ниже показано одно примерное решение, где $b1$ является некоторой скоростью передачи битов, например, 48 Кбит/с, где TCX_{20} указывает то, что кадр кодируется с

использованием одного длинного блока, где TCX_10 указывает то, что кадр кодируется с использованием 2, 3, 4 или более коротких блоков, где решение по выбору TCX_20/TCX_10 основано на выводе детектора переходных частей, описанного выше; tempFlatness является показателем временной равномерности, как задано в (6), maxEnergyChange является максимальным изменением энергии, как задано в (7). Условие "norm_corr(curr)>1,2- T_{int}/L " также может записываться как "(1,2-norm_corr(curr))*L< T_{int} ".

```

enableLTP =
  (скорость передачи битов < b1 && tcxMode == TCX_20 && (norm_corr(curr) * norm_corr(prev)) > 0,25
  && tempFlatness < 3.5) ||
  (скорость передачи битов >= b1 && tcxMode == TCX_10 && макс. (norm_corr(curr), norm_corr(prev)) >
  0,5 && maxEnergyChange < 3.5) ||
  (скорость передачи битов >= b1 && norm_corr(curr) > 0,44 && norm_corr(curr) > 1.2- $T_{int}/L$ ) ||
  (скорость передачи битов >= b1 && tcxMode == TCX_20 && norm_corr(curr) > 0,44 &&
  (tempFlatness < 6.0 || (tempFlatness < 7,0 && maxEnergyChange < 22.0)));
  (скорость передачи битов >= b1 && tcxMode == TCX_20 && norm_corr > 0,44 &&

```

Из вышеприведенных примеров очевидно то, что обнаружение переходных частей влияет на то, какой механизм решения для долговременного прогнозирования используется, и какая часть сигнала используется для измерений, используемых в решении, а не на то, что оно непосредственно инициирует деактивацию фильтра долговременного прогнозирования.

Временные показатели, используемые для решения длины преобразования, могут абсолютно отличаться от временных показателей, используемых для решения по выбору ЛТР-фильтра, либо они могут перекрываться или быть совершенно идентичными, но вычисляться в различных областях. Для низкотоновых сигналов, обнаружение переходных частей может полностью игнорироваться, если достигнуто пороговое значение для нормализованной корреляции, которое зависит от запаздывания основного тона.

Технология для удаления возможных разрывов

Далее описывается возможная технология для удаления разрывов, вызываемых посредством покадрового применения линейного фильтра $H(z)$. Линейный фильтр может представлять собой описанный ЛТР-фильтр. Линейный фильтр может представлять собой фильтр FIR (с конечной импульсной характеристикой) или фильтр IIR (с бесконечной импульсной характеристикой). Предложенный подход не фильтрует часть текущего кадра с помощью параметров фильтрации предыдущего кадра и за счет этого не допускает возможных проблем известных подходов. Предложенный подход использует LPC-фильтр, чтобы удалять разрыв. Этот LPC-фильтр оценивается для аудиосигнала (фильтрованного или нет посредством линейного независимого от времени фильтра $H(z)$) и в силу этого представляет собой хорошую модель спектральной формы аудиосигнала (фильтрованного или нет посредством $H(z)$). LPC-фильтр затем используется таким образом, что спектральная форма аудиосигнала маскирует разрыв.

LPC-фильтр может оцениваться по-разному. Он может оцениваться, например, с использованием аудиосигнала (текущего и/или предыдущего кадра) и алгоритма Левинсона-Дурбина. Он также может вычисляться для сигнала предыдущего фильтрованного кадра, с использованием алгоритма Левинсона-Дурбина.

Если $H(z)$ используется в аудиокодеке, и аудиокодек уже использует LPC-фильтр (квантованный или нет) для того, чтобы, например, придавать определенную форму шуму квантования в аудиокодеке на основе преобразования, то этот LPC-фильтр может

быть непосредственно использован для сглаживания разрыва, без дополнительной сложности, требуемой для того, чтобы оценивать новый LPC-фильтр.

Ниже описывается обработка текущего кадра для случая FIR-фильтра и случая IIR-фильтра. Предыдущий кадр предполагается как уже обработанный.

5 Случай FIR-фильтра:

Фильтрация текущего кадра с помощью параметров фильтрации текущего кадра, формируя фильтрованный текущий кадр.

Рассмотрение LPC-фильтра (квантованного или нет) с порядком M , оцененным для аудиосигнала (фильтрованного или нет).

10 M последних выборок предыдущего кадра фильтруются с помощью фильтра $H(z)$ и коэффициентов текущего кадра, формируя первую часть фильтрованного сигнала.

M последних выборок фильтрованного предыдущего кадра затем вычитаются из первой части фильтрованного сигнала, формируя вторую часть фильтрованного сигнала.

15 Нулевая импульсная характеристика (ZIR) LPC-фильтра затем формируется посредством фильтрации кадра нулевых выборок с помощью LPC-фильтра и начальных состояний, равных второй части фильтрованного сигнала.

ZIR необязательно может кодироваться со взвешиванием таким образом, что его амплитуда быстрее достигает 0.

20 Начальная часть ZIR вычитается из соответствующей начальной части фильтрованного текущего кадра.

Случай IIR-фильтра:

Рассмотрение LPC-фильтра (квантованного или нет) с порядком M , оцененным для аудиосигнала (фильтрованного или нет).

25 M последних выборок предыдущего кадра фильтруются с помощью фильтра $H(z)$ и коэффициентов текущего кадра, формируя первую часть фильтрованного сигнала.

M последних выборок фильтрованного предыдущего кадра затем вычитаются из первой части фильтрованного сигнала, формируя вторую часть фильтрованного сигнала.

30 Нулевая импульсная характеристика (ZIR) LPC-фильтра затем формируется посредством фильтрации кадра нулевых выборок с помощью LPC-фильтра и начальных состояний, равных второй части фильтрованного сигнала.

ZIR необязательно может кодироваться со взвешиванием таким образом, что его амплитуда быстрее достигает 0.

Начальная часть текущего кадра затем обрабатывается последовательно выборочно начиная с первой выборки текущего кадра.

35 Выборка фильтруется с помощью фильтра $H(z)$ и параметров текущего кадра, формируя первую фильтрованную выборку.

Соответствующая выборка ZIR затем вычитается из первой фильтрованной выборки, формируя соответствующую выборку фильтрованного текущего кадра.

Переход к следующей выборке.

40 Повтор 9-12 до тех пор, пока не будет обработана последняя выборка начальной части текущего кадра.

Фильтрация оставшихся выборок текущего кадра с помощью параметров фильтрации текущего кадра.

45 Соответственно, варианты осуществления изобретения обеспечивают оценку сегментальных SNR и выбор надлежащего алгоритма кодирования простым и точным способом. В частности, варианты осуществления изобретения обеспечивают выбор с разомкнутым контуром надлежащего алгоритма кодирования, при этом несоответствующий выбор алгоритма кодирования в случае аудиосигнала, имеющего

гармоники, не допускается.

В вышеописанных вариантах осуществления, сегментальные SNR оцениваются посредством вычисления среднего SNR, оцененных для соответствующих субкадров. В альтернативных вариантах осуществления, может оцениваться SNR всего кадра без

5 разделения кадра на субкадры.

Варианты осуществления изобретения обеспечивают существенное уменьшение времени вычислений, по сравнению с выбором с замкнутым контуром, поскольку ряд этапов, требуемых при выборе с замкнутым контуром, опускаются.

Соответственно, большое число этапов и время вычислений, ассоциированное с

10 ними, может экономиться посредством изобретаемого подхода при одновременном обеспечении выбора надлежащего алгоритма кодирования с хорошей производительностью.

Хотя некоторые аспекты описаны в контексте устройства, очевидно, что эти аспекты также представляют описание соответствующего способа, при этом блок или устройство

15 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.

Варианты осуществления устройств, описанных в данном документе, и их признаков могут реализовываться посредством компьютера, одного или более процессоров,

20 одного или более микропроцессоров, программируемых пользователем вентильных матриц (FPGA), специализированных интегральных схем (ASIC) и т.п. либо комбинаций вышеозначенного, которые сконфигурированы или запрограммированы с возможностью предоставлять описанные функциональности.

Некоторые или все этапы способа могут быть выполнены посредством (или с

25 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

30 изобретения могут быть реализованы в аппаратных средствах или в программном обеспечении. Реализация может выполняться с использованием с использованием энергонезависимого носителя хранения данных, такого как цифровой носитель хранения данных, например, гибкий диск, DVD, Blu-Ray, CD, ROM, PROM и EPROM, EEPROM или флэш-память, имеющего сохраненные электронночитаемые управляющие сигналы,

35 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

40 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

45 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 выполнена с возможностью передачи через соединение для передачи данных, например, через Интернет.

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

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 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.

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 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 и изменения компоновок и подробностей, описанных в данном документе, должны быть очевидными для специалистов в данной области техники. Следовательно, они подразумеваются как ограниченные только объемом нижеприведенной формулы изобретения, а не конкретными деталями, представленными посредством описания и пояснения вариантов осуществления в данном документе.

(57) Формула изобретения

1. Устройство (10) для выбора одного из первого алгоритма кодирования, имеющего первую характеристику, и второго алгоритма кодирования, имеющего вторую характеристику, для кодирования части аудиосигнала (40), чтобы получить кодированную версию части аудиосигнала (40), содержащее:

- фильтр долговременного прогнозирования, выполненный с возможностью принимать аудиосигнал, уменьшать амплитуду гармоник в аудиосигнале и выводить фильтрованную версию аудиосигнала;

- первый модуль (12) оценки для использования фильтрованной версии аудиосигнала при оценке SNR (отношения "сигнал-шум") или сегментального SNR части аудиосигнала в качестве первого показателя качества для части аудиосигнала, причем первый показатель качества ассоциирован с первым алгоритмом кодирования, при этом оценка упомянутого первого показателя качества содержит выполнение аппроксимации первого алгоритма кодирования для того, чтобы получить оценку искажения первого алгоритма кодирования и оценивать первый показатель качества на основе части аудиосигнала, и оценку искажения первого алгоритма кодирования без фактического кодирования и декодирования части аудиосигнала с использованием первого алгоритма кодирования;

- второй модуль (14) оценки для оценки SNR или сегментального SNR в качестве второго показателя качества для части аудиосигнала, причем второй показатель качества ассоциирован со вторым алгоритмом кодирования, при этом оценка упомянутого второго показателя качества содержит выполнение аппроксимации второго алгоритма кодирования для того, чтобы получить оценку искажения второго алгоритма кодирования и оценивать второй показатель качества с использованием части аудиосигнала, и оценку искажения второго алгоритма кодирования без фактического кодирования и декодирования части аудиосигнала с использованием второго алгоритма кодирования; и

- контроллер (16) для выбора первого алгоритма кодирования или второго алгоритма кодирования на основе сравнения между первым показателем качества и вторым показателем качества,

- при этом первый алгоритм кодирования представляет собой алгоритм кодирования с преобразованием, алгоритм кодирования на основе MDCT (модифицированного дискретного косинусного преобразования) или алгоритм кодирования на основе TCX (возбуждения по кодированию с преобразованием), при этом второй алгоритм кодирования представляет собой алгоритм кодирования на основе CELP (линейного прогнозирования с возбуждением по коду) или алгоритм кодирования на основе ACELP (линейного прогнозирования с возбуждением по алгебраическому коду).

2. Устройство (10) по п. 1, в котором передаточная функция фильтра долговременного прогнозирования содержит целочисленную часть запаздывания основного тона и многоотводный фильтр в зависимости от дробной части запаздывания основного тона.

3. Устройство (10) по п. 1, в котором фильтр долговременного прогнозирования имеет передаточную функцию

$$P(z) = 1 - \beta g B(z, T_{fr}) z^{-T_{int}},$$

где T_{int} и T_{fr} являются целочисленной и дробной частью запаздывания основного тона, g усиление, β является весовым коэффициентом, и $B(z, T_{fr})$ является FIR-фильтром нижних частот, коэффициенты которого зависят от дробной части основного тона.

4. Устройство по п. 1, дополнительно содержащее модуль деактивации для деактивации фильтра на основе комбинации одного или более показателей гармонического характера и/или одного или более показателей временной структуры.

5. Устройство по п. 4, в котором один или более показателей гармонического характера содержат по меньшей мере одно из нормализованной корреляции или усиления для прогнозирования, при этом один или более показателей временной структуры содержат по меньшей мере одно из показателя временной равномерности и изменения энергии.

6. Устройство по п. 1, в котором фильтр применяется к аудиосигналу на покадровой основе, причем упомянутое устройство дополнительно содержит модуль для удаления разрывов в аудиосигнале, вызываемых фильтром.

7. Устройство (10) по п. 1, в котором первый и второй модули оценки выполнены с возможностью оценивать SNR или сегментальное SNR части взвешенной версии аудиосигнала.

8. Устройство (10) по п. 1, в котором первый модуль (12) оценки выполнен с возможностью определять оцененное искажение квантователя, которое квантователь, используемый в первом алгоритме кодирования, должен вводить при квантовании части аудиосигнала, и оценивать первый показатель качества на основе энергии части взвешенной версии аудиосигнала и оцененного искажения квантователя, при этом первый модуль (12) оценки выполнен с возможностью оценивать глобальное усиление для части аудиосигнала таким образом, что часть аудиосигнала должна формировать данную целевую скорость передачи битов при кодировании с помощью квантователя и энтропийного кодера, используемых в первом алгоритме кодирования, при этом первый модуль (12) оценки дополнительно выполнен с возможностью определять оцененное искажение квантователя на основе оцененного глобального усиления.

9. Устройство (10) по п. 1, в котором второй модуль (14) оценки выполнен с возможностью определять оцененное искажение адаптивной таблицы кодирования, которое адаптивная таблица кодирования, используемая во втором алгоритме кодирования, должна вводить при использовании адаптивной таблицы кодирования для того, чтобы кодировать часть аудиосигнала, при этом второй модуль (14) оценки выполнен с возможностью оценивать второй показатель качества на основе энергии части взвешенной версии аудиосигнала и оцененного искажения адаптивной таблицы кодирования, при этом для каждой из множества подчастей части аудиосигнала второй модуль (14) оценки выполнен с возможностью аппроксимировать адаптивную таблицу кодирования на основе версии подчасти взвешенного аудиосигнала, сдвинутой в прошлое посредством запаздывания основного тона, определенного в каскаде предварительной обработки, чтобы оценивать усиление адаптивной таблицы кодирования таким образом, что ошибка между подчастью части взвешенного аудиосигнала и аппроксимированной адаптивной таблицей кодирования минимизируется, и определять оцененное искажение адаптивной таблицы кодирования на основе энергии ошибки между подчастью части взвешенного аудиосигнала и аппроксимированной адаптивной таблицей кодирования, масштабированной посредством усиления адаптивной таблицы кодирования.

10. Устройство (10) по п. 9, в котором второй модуль (14) оценки дополнительно выполнен с возможностью уменьшать оцененное искажение адаптивной таблицы кодирования, определенное для каждой подчасти части аудиосигнала, посредством постоянного множителя.

11. Устройство (10) по п. 1, в котором второй модуль (14) оценки выполнен с возможностью определять оцененное искажение адаптивной таблицы кодирования, которое адаптивная таблица кодирования, используемая во втором алгоритме кодирования, должна вводить при использовании адаптивной таблицы кодирования

для того, чтобы кодировать часть аудиосигнала, при этом второй модуль (14) оценки выполнен с возможностью оценивать второй показатель качества на основе энергии части взвешенной версии аудиосигнала и оцененного искажения адаптивной таблицы кодирования, при этом второй модуль (14) оценки выполнен с возможностью
 5 аппроксимировать адаптивную таблицу кодирования на основе версии части взвешенного аудиосигнала, сдвинутой в прошлое посредством запаздывания основного тона, определенного в каскаде предварительной обработки, чтобы оценивать усиление адаптивной таблицы кодирования таким образом, что ошибка между частью взвешенного аудиосигнала и аппроксимированной адаптивной таблицей кодирования
 10 минимизируется, и определять оцененное искажение адаптивной таблицы кодирования на основе энергии ошибки между частью взвешенного аудиосигнала и аппроксимированной адаптивной таблицей кодирования, масштабированной посредством усиления адаптивной таблицы кодирования.

12. Устройство (20) для кодирования части аудиосигнала, содержащее устройство
 15 (10) по одному из пп. 1-11, первый каскад (26) кодера для выполнения первого алгоритма кодирования и второй каскад (28) кодера для выполнения второго алгоритма кодирования, при этом устройство (20) для кодирования выполнено с возможностью кодировать часть аудиосигнала с использованием первого алгоритма кодирования или второго алгоритма кодирования в зависимости от выбора посредством контроллера
 20 (16).

13. Система для кодирования и декодирования, содержащая устройство (20) для кодирования по п. 12 и декодер, выполненный с возможностью принимать кодированную версию части аудиосигнала и индикатор относительно алгоритма, используемого для того, чтобы кодировать часть аудиосигнала и декодировать
 25 кодированную версию части аудиосигнала с использованием указываемого алгоритма.

14. Способ выбора одного из первого алгоритма кодирования, имеющего первую характеристику, и второго алгоритма кодирования, имеющего вторую характеристику, для кодирования части аудиосигнала, чтобы получать кодированную версию части аудиосигнала, содержащий этапы, на которых:

30 - фильтруют аудиосигнал с использованием фильтра долговременного прогнозирования, чтобы уменьшать амплитуду гармоник в аудиосигнале и выводить фильтрованную версию аудиосигнала;

- используют фильтрованную версию аудиосигнала при оценке SNR или сегментальное SNR части аудиосигнала в качестве первого показателя качества для части аудиосигнала, причем первый показатель качества ассоциирован с первым алгоритмом кодирования, при этом оценка упомянутого первого показателя качества содержит этап, на котором
 35 выполняют аппроксимацию первого алгоритма кодирования для того, чтобы получать оценку искажения первого алгоритма кодирования и оценивать первый показатель качества на основе части первого аудиосигнала, и оценивают искажение первого
 40 алгоритма кодирования без фактического кодирования и декодирования части аудиосигнала с использованием первого алгоритма кодирования;

- оценивают SNR или сегментальное SNR в качестве второго показателя качества для части аудиосигнала, причем второй показатель качества ассоциирован со вторым алгоритмом кодирования, при этом оценка упомянутого второго показателя качества
 45 содержит этап, на котором выполняют аппроксимацию второго алгоритма кодирования для того, чтобы получать оценку искажения второго алгоритма кодирования и оценивать второй показатель качества с использованием части аудиосигнала, и оценивают искажение второго алгоритма кодирования без фактического кодирования и

декодирования части аудиосигнала с использованием второго алгоритма кодирования;
и

- выбирают первый алгоритм кодирования или второй алгоритм кодирования на основе сравнения между первым показателем качества и вторым показателем качества,

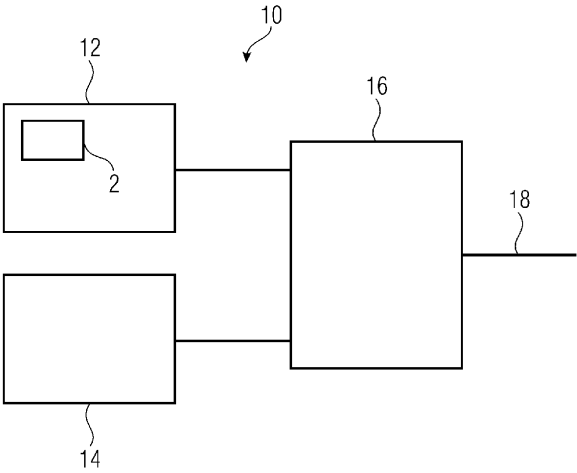
- при этом первый алгоритм кодирования представляет собой алгоритм кодирования с преобразованием, алгоритм кодирования на основе MDCT (модифицированного дискретного косинусного преобразования) или алгоритм кодирования на основе TCX (возбуждения по кодированию с преобразованием), при этом второй алгоритм кодирования представляет собой алгоритм кодирования на основе CELP (линейного прогнозирования с возбуждением по коду) или алгоритм кодирования на основе ACELP (линейного прогнозирования с возбуждением по алгебраическому коду).

15. Машиночитаемый носитель данных, содержащий компьютерную программу, имеющую программный код для осуществления, при выполнении на компьютере, способа по одному из п. 14.

1

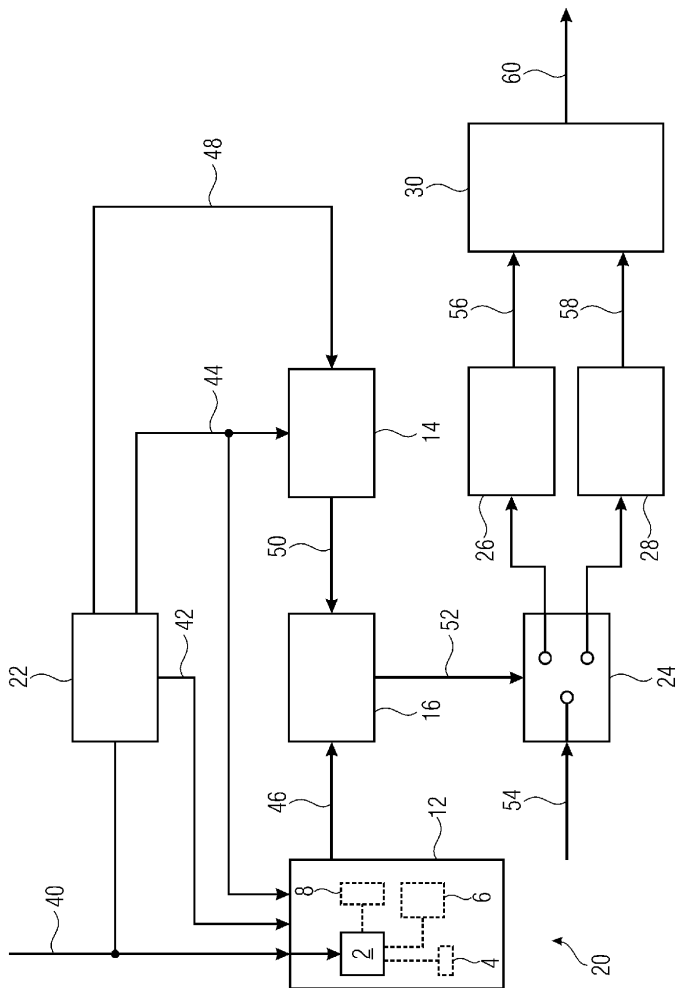
529771

1/4

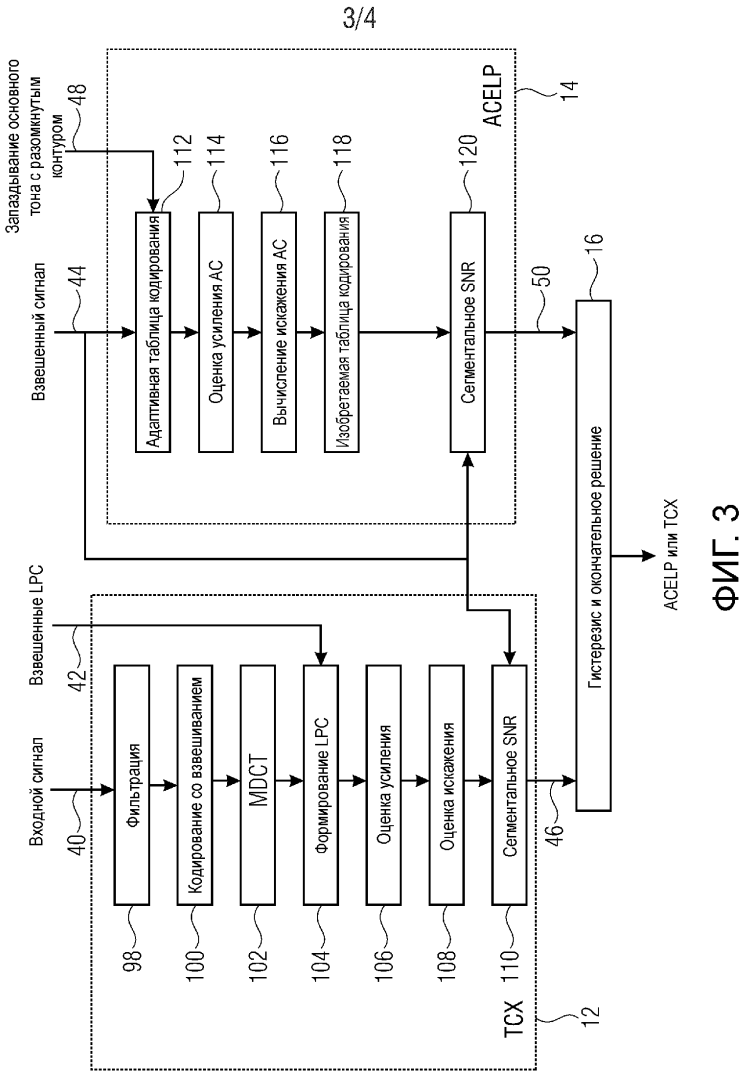


ФИГ. 1

2



ФИГ. 2



4/4

$$\text{SNR} = 10 \log_{10} \frac{\sum_{i=1}^N x^2(i)}{\sum_{i=1}^N (x(i) - y(i))^2}$$

ФИГ. 4А

$$\text{SNR}_{\text{seg}} = \frac{10}{M} \sum_{m=0}^{M-1} \log_{10} \sum_{i=Nm}^{Nm-N-1} \left(\frac{\sum_{i=1}^N x^2(i)}{\sum_{i=1}^N (x(i) - y(i))^2} \right)$$

ФИГ. 4В