



ФЕДЕРАЛЬНАЯ СЛУЖБА
ПО ИНТЕЛЛЕКТУАЛЬНОЙ СОБСТВЕННОСТИ

(12) ОПИСАНИЕ ИЗОБРЕТЕНИЯ К ПАТЕНТУ

(52) СПК

G10L 19/032 (2006.01); G10L 19/083 (2006.01); G10L 19/12 (2006.01); G10L 19/20 (2006.01)

(21)(22) Заявка: 2016119010, 10.10.2014

(24) Дата начала отсчета срока действия патента:
10.10.2014

Дата регистрации:
02.03.2018

Приоритет(ы):

(30) Конвенционный приоритет:
18.10.2013 ЕР ЕР13189392.7;
28.07.2014 ЕР 14178788.7

(43) Дата публикации заявки: 23.11.2017 Бюл. № 33

(45) Опубликовано: 02.03.2018 Бюл. № 7

(85) Дата начала рассмотрения заявки РСТ на
национальной фазе: 18.05.2016

(86) Заявка РСТ:
ЕР 2014/071767 (10.10.2014)

(87) Публикация заявки РСТ:
WO 2015/055531 (23.04.2015)

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

(72) Автор(ы):

ФУКС Гийом (DE),
МУЛЬТРУС Маркус (DE),
РАВЕЛЛИ Эммануэль (DE),
ШНЕЛЛЬ Маркус (DE)

(73) Патентообладатель(и):

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

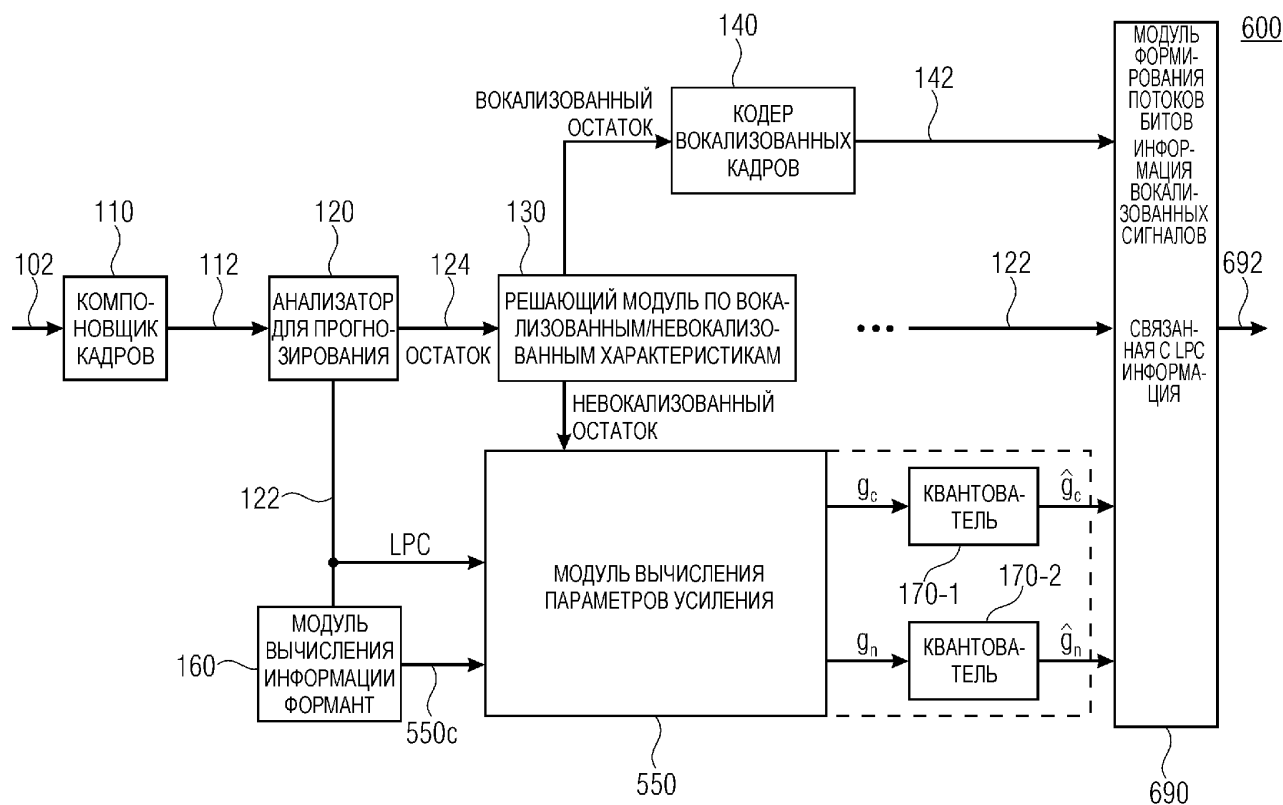
(56) Список документов, цитированных в отчете
о поиске: US 6,611,800 B1, 26.08.2003. US
5,732,389 A, 24.03.1998. ЕР 0 967 594 A1,
29.12.1999. US 2005/0010402 A1, 13.01.2005. RU
2012 130 472 A, 10.09.2013.

(54) ПРИНЦИП ДЛЯ КОДИРОВАНИЯ АУДИОСИГНАЛА И ДЕКОДИРОВАНИЯ АУДИОСИГНАЛА С ИСПОЛЬЗОВАНИЕМ ИНФОРМАЦИИ ФОРМИРОВАНИЯ РЕЧЕВОГО СПЕКТРА

(57) Реферат:

Изобретение относится к кодированию аудиосигнала. Технический результат изобретения заключается в возможности поддерживать качество речи при уменьшении скорости передачи битов. Кодер содержит анализатор, выполненный с возможностью извлечения коэффициентов прогнозирования и остаточного сигнала из кадра аудиосигнала. Кодер содержит модуль вычисления информации формант для вычисления информации формирования речевого спектра из

коэффициентов прогнозирования, модуль вычисления параметров усиления, выполненный с возможностью вычисления параметра усиления из невокализованного остаточного сигнала и информации формирования спектра, и модуль формирования потоков битов, выполненный с возможностью формирования выходного сигнала на основе информации, связанной с вокализованным кадром сигнала, параметром усиления или параметром квантованного



ФИГ. 6



FEDERAL SERVICE
FOR INTELLECTUAL PROPERTY

(12) **ABSTRACT OF INVENTION**

(52) CPC

G10L 19/032 (2006.01); G10L 19/083 (2006.01); G10L 19/12 (2006.01); G10L 19/20 (2006.01)(21)(22) Application: **2016119010, 10.10.2014**(24) Effective date for property rights:
10.10.2014Registration date:
02.03.2018

Priority:

(30) Convention priority:
18.10.2013 EP EP13189392.7;
28.07.2014 EP 14178788.7(43) Application published: **23.11.2017 Bull. № 33**(45) Date of publication: **02.03.2018 Bull. № 7**(85) Commencement of national phase: **18.05.2016**(86) PCT application:
EP 2014/071767 (10.10.2014)(87) PCT publication:
WO 2015/055531 (23.04.2015)Mail address:
129090, Moskva, ul. B.Spasskaya, 25, stroenie 3,
OOO "Yuridicheskaya firma Gorodisskij i Partnery"

(72) Inventor(s):

FUKS Gijom (DE),
MULTRUS Markus (DE),
RAVELLI Emmanuel (DE),
SHNELL Markus (DE)

(73) Proprietor(s):

FRAUNKHOFER-GEZELLSCHAFT TSUR
FERDERUNG DER ANGEVANDTEN
FORSHUNG E.F. (DE)(54) **PRINCIPLE FOR CODING AUDIO SIGNAL AND DECODING AUDIO SIGNAL USING INFORMATION FOR GENERATING SPEECH SPECTRUM**

(57) Abstract:

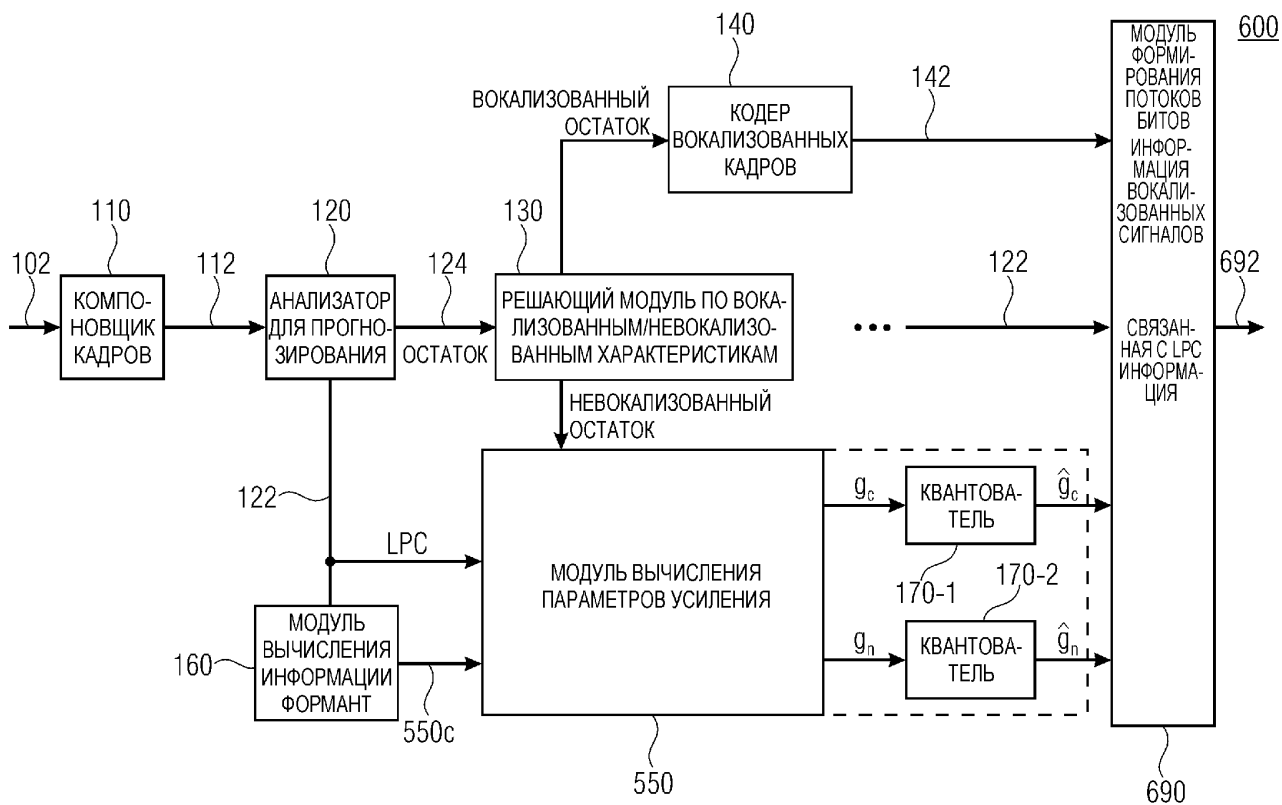
FIELD: physics.

SUBSTANCE: encoder comprises an analyzer configured to extract the prediction coefficients and the residual signal from the audio signal frame. The encoder contains a module for computing formant information to calculate information of generating a speech spectrum from the prediction coefficients, a module of calculating the gain parameters, configured to calculate the gain parameter of the unvoiced residual signal and

information of generating the spectrum, and a module of generating the streams of bits, configured to generate the output signal based on information associated with a vocalized signal frame, a gain parameter, or a quantized gain parameter and predictor coefficients.

EFFECT: ability to maintain the voice quality while reducing the bit rate.

18 cl, 17 dwg



ФИГ. 6

Настоящее изобретение относится к кодерам для кодирования аудиосигнала, в частности, речевого аудиосигнала. Настоящее изобретение также относится к декодерам и способам для декодирования кодированного аудиосигнала. Настоящее изобретение дополнительно относится к кодированным аудиосигналам и к усовершенствованному кодированию невокализованной речи на низких скоростях передачи битов.

На низкой скорости передачи битов, кодирование речи может извлекать выгоду из специальной обработки для невокализованных кадров, чтобы поддерживать качество речи при уменьшении скорости передачи битов. Невокализованные кадры могут перцепционно моделироваться в качестве случайного возбуждения, которое формируется в частотной и временной области. Поскольку форма сигнала и возбуждение выглядит и звучит почти идентично с белым гауссовым шумом, его кодирование на основе формы сигналов может ослабляться и заменяться посредством синтетически сформированного белого шума. В таком случае кодирование должно состоять из кодирования форм во временной и частотной области сигнала.

Фиг. 16 показывает принципиальную блок-схему схемы параметрического невокализованного кодирования. Синтезирующий фильтр 1202 выполнен с возможностью моделирования речевого тракта и параметризован посредством параметров LPC (линейного прогнозирующего кодирования). Из извлеченного LPC-фильтра, содержащего функцию $A(z)$ фильтра, перцепционный взвешивающий фильтр может извлекаться посредством взвешивания LPC-коэффициентов. Перцепционный фильтр $fw(n)$ обычно имеет передаточную функцию следующей формы:

$$Ffw(z) = \frac{A(z)}{A(z/w)},$$

где w ниже 1. Параметр g_n усиления вычисляется для получения синтезированной энергии, совпадающей с исходной энергией в перцепционной области, согласно следующему:

$$g_n = \sqrt{\frac{\sum_{n=0}^{Ls} sw^2(n)}{\sum_{n=0}^{Ls} nw^2(n)}},$$

где $sw(n)$ и $nw(n)$ являются входным сигналом и сформированным шумом, соответственно, фильтруемыми посредством перцепционного фильтра $fw(n)$. Усиление g_n вычисляется для каждого субкадра размера Ls . Например, аудиосигнал может быть разделен на кадры с длиной в 20 мс. Каждый кадр может подразделяться на субкадры, например, на четыре субкадра, каждый из которых имеет длину в 5 мс.

Схема кодирования на основе линейного прогнозирования с возбуждением по коду (CELP) широко используется в речевой связи и представляет собой очень эффективный способ кодирования речи. Она обеспечивает более естественное качество речи, чем параметрическое кодирование, но она также запрашивает более высокие скорости. CELP синтезирует аудиосигнал посредством передачи в линейный прогнозирующий фильтр, называемый синтезирующим LPC-фильтром, который может содержать форму $1/A(z)$, суммы двух возбуждений. Одно возбуждение исходит из декодированного прошлого, что называется адаптивной таблицей кодирования. Другая доля исходит из изобретаемой таблицы кодирования, заполненной посредством фиксированных кодов. Тем не менее, на низких скоростях передачи битов изобретаемая таблица кодирования недостаточно заполняется для эффективного моделирования точной структуры речи или шумоподобного возбуждения невокализованной речи. Следовательно, перцепционное качество ухудшается, в частности, невокализованные кадры, которые

в таком случае звучат трескуче и неестественно.

Для уменьшения артефактов кодирования на низких скоростях передачи битов, уже предложены различные решения. В G.718[1] и в [2], коды изобретаемой таблицы кодирования адаптивно и спектрально формируются посредством улучшения спектральных областей, соответствующих формантам текущего кадра. Позиции и формы формант могут быть выведены непосредственно из LPC-коэффициентов, причем коэффициенты уже доступны на сторонах кодера и декодера. Улучшение формант кодов $c(n)$ выполняется посредством простой фильтрации согласно следующему:

$$c(n) * fe(n),$$

где $*$ обозначает оператор свертки, и где $fe(n)$ является импульсной характеристикой фильтра передаточной функции следующим образом:

$$Ffe(z) = \frac{A(z/w1)}{A(z/w2)},$$

где $w1$ и $w2$ являются двумя весовыми константами, подчеркивающими более или менее формантную структуру передаточной функции $Ffe(z)$. Результирующие коды определенной формы наследуют характеристику речевого сигнала, и синтезированный сигнал звучит чище.

В CELP, также обычно добавляется спектральный наклон в декодер изобретаемой таблицы кодирования. Это выполняется посредством фильтрации кодов с помощью следующего фильтра:

$$Ft(z) = 1 - \beta z^{-1}$$

Коэффициент β обычно связан с вокализацией предыдущего кадра и зависит, т.е. он варьируется. Вокализация может оцениваться из доли энергии из адаптивной таблицы кодирования. Если предыдущий кадр является вокализованным, предполагается, что текущий кадр также является вокализованным, и что коды должны иметь большую энергию на низких частотах, т.е. должны показывать отрицательный наклон. Наоборот, добавленный спектральный наклон является положительным для невокализованных кадров, и большая энергия должна распределяться в направлении высоких частот.

Использование формирования спектра для улучшения речи и уменьшения уровня шума выхода декодера является обычной практикой. Так называемое улучшение формант в качестве постфильтрации состоит из адаптивной постфильтрации, для которой коэффициенты извлекаются из LPC-параметров декодера. Постфильтр похож на постфильтр $(fe(n))$, используемый для формирования изобретаемого возбуждения в определенных CELP-кодерах, как пояснено выше. Тем не менее, в этом случае, постфильтрация применяется только на конце процесса декодера, а не на стороне кодера.

В традиционном CELP (CELP – линейное прогнозирование с возбуждением по таблицам кодирования), форма частоты моделируется посредством синтезирующего фильтра на основе LP (линейного прогнозирования), в то время как форма во временной области может быть аппроксимирована посредством усиления при возбуждении, отправленного в каждый субкадр, хотя долговременное прогнозирование (LTP) и изобретаемая таблица кодирования обычно не подходят для моделирования шумоподобного возбуждения невокализованных кадров. CELP требует относительно высокой скорости передачи битов для достижения хорошего качества невокализованной речи.

Определение вокализованных или невокализованных характеристик может быть связано с сегментацией речи на части и ассоциированием каждой из них с различной исходной моделью речи. Исходные модели, когда они используются в схеме CELP-

кодирования речи, основываются на адаптивном гармоническом возбуждении, моделирующем поток воздуха, выходящий из голосовой щели, и резонансном фильтре, моделирующем речевой тракт, возбужденный посредством сформированного потока воздуха. Такие модели могут предоставлять хорошие результаты для фонем, таких как
 5 вокалы, но могут приводить к некорректному моделированию для речевых частей, которые не формируются посредством голосовой щели, в частности, когда голосовые связки не вибрируют, к примеру, невокализованных фонем "s" или "f".

С другой стороны, параметрические речевые кодеры также называются вокодерами и приспособливают одну исходную модель для невокализованных кадров. Она позволяет
 10 достигать очень низких скоростей передачи битов при достижении так называемого синтетического качества, не настолько естественного, насколько качество, обеспечиваемое посредством схем CELP-кодирования на гораздо более высоких скоростях.

Таким образом, имеется потребность в улучшении аудиосигналов.

15 Цель настоящего изобретения заключается в том, чтобы повышать качество звука на низких скоростях передачи битов и/или уменьшать скорости передачи битов для хорошего качества звука.

Это цель достигается посредством кодера, декодера, кодированного аудиосигнала и способов согласно независимым пунктам формулы изобретения.

20 Авторы изобретения выяснили, что в первом аспекте, качество декодированного аудиосигнала, связанного с невокализованным кадром аудиосигнала, может увеличиваться, т.е. повышаться, посредством определения информации формирования речевого спектра таким образом, что информация параметров усиления для усиления сигналов может извлекаться из информации формирования речевого спектра. Кроме
 25 того информация формирования речевого спектра может использоваться для спектрального формирования декодированного сигнала. Частотные области, содержащие более высокую важность для речи, например, низкие частоты ниже 4 кГц, в силу этого могут обрабатываться таким образом, что они содержат меньше ошибок.

Авторы изобретения дополнительно выяснили, что во втором аспекте, посредством
 30 формирования первого сигнала возбуждения из детерминированной таблицы кодирования для кадра или субкадра (части) синтезированного сигнала и посредством формирования второго сигнала возбуждения из шумоподобного сигнала для кадра или субкадра синтезированного сигнала и посредством комбинирования первого сигнала возбуждения и второго сигнала возбуждения для формирования

35 комбинированного сигнала возбуждения качество звука синтезированного сигнала может увеличиваться, т.е. повышаться. В частности, для частей аудиосигнала, содержащего речевой сигнал с фоновым шумом, качество звука может повышаться посредством добавления шумоподобных сигналов. Параметр усиления для необязательного усиления первого сигнала возбуждения может определяться в кодере,
 40 и информация, связанная с ним, может передаваться с кодированным аудиосигналом.

Альтернативно или помимо этого, улучшение синтезированного аудиосигнала может быть, по меньшей мере, частично использовано для уменьшения скоростей передачи битов для кодирования аудиосигнала.

Кодер согласно первому аспекту содержит анализатор, выполненный с возможностью
 45 извлечения коэффициентов прогнозирования и остаточного сигнала из кадра аудиосигнала. Кодер дополнительно содержит модуль вычисления информации формант, выполненный с возможностью вычисления информации формирования речевого спектра из коэффициентов прогнозирования. Кодер дополнительно содержит модуль вычисления

параметров усиления, выполненный с возможностью вычисления параметра усиления из невокализованного остаточного сигнала и информации формирования спектра, и модуль формирования потоков битов, выполненный с возможностью формирования выходного сигнала на основе информации, связанной с вокализованным кадром сигнала, параметром усиления или параметром квантованного усиления и коэффициентами прогнозирования.

Дополнительные варианты осуществления первого аспекта предоставляют кодированный аудиосигнал, содержащий информацию коэффициентов прогнозирования для вокализованного кадра и невокализованного кадра аудиосигнала, дополнительную информацию, связанную с вокализованным кадром сигнала, и параметр усиления либо параметр квантованного усиления для невокализованного кадра. Это дает возможность эффективной передачи речевой информации, чтобы обеспечивать декодирование кодированного аудиосигнала с тем, чтобы получать синтезированный (восстановленный) сигнал с высоким качеством звука.

Дополнительные варианты осуществления первого аспекта предоставляют декодер для декодирования принимаемого сигнала, содержащего коэффициенты прогнозирования. Декодер содержит модуль вычисления информации формант, генератор шума, формирователь и синтезатор. Модуль вычисления информации формант выполнен с возможностью вычисления информации формирования речевого спектра из коэффициентов прогнозирования. Генератор шума выполнен с возможностью формирования шумоподобного сигнала для декодирования. Формирователь выполнен с возможностью придания определенной формы спектру шумоподобного сигнала для декодирования или его усиленному представлению с использованием информации формирования спектра, чтобы получать шумоподобный сигнал для декодирования определенной формы. Синтезатор выполнен с возможностью синтеза синтезированного сигнала из усиленного шумоподобного сигнала для кодирования определенной формы и коэффициентов прогнозирования.

Дополнительные варианты осуществления первого аспекта относятся к способу для кодирования аудиосигнала, к способу для декодирования принимаемого аудиосигнала и к компьютерной программе.

Варианты осуществления второго аспекта предоставляют кодер для кодирования аудиосигнала. Кодер содержит анализатор, выполненный с возможностью извлечения коэффициентов прогнозирования и остаточного сигнала из невокализованного кадра аудиосигнала. Кодер дополнительно содержит модуль вычисления параметров усиления, выполненный с возможностью вычисления информации первых параметров усиления для задания первого сигнала возбуждения, связанного с детерминированной таблицей кодирования, и вычисления информации вторых параметров усиления для задания второго сигнала возбуждения, связанного с шумоподобным сигналом для невокализованного кадра. Кодер дополнительно содержит модуль формирования потоков битов, выполненный с возможностью формирования выходного сигнала на основе информации, связанной с вокализованным кадром сигнала, информации первых параметров усиления и информации вторых параметров усиления.

Дополнительные варианты осуществления второго аспекта предоставляют декодер для декодирования принимаемого аудиосигнала, содержащего информацию, связанную с коэффициентами прогнозирования. Декодер содержит генератор первых сигналов, выполненный с возможностью формирования первого сигнала возбуждения из детерминированной таблицы кодирования для части синтезированного сигнала. Декодер дополнительно содержит генератор вторых сигналов, выполненный с возможностью

формирования второго сигнала возбуждения из шумоподобного сигнала для части синтезированного сигнала. Декодер дополнительно содержит модуль комбинирования и синтезатор, при этом модуль комбинирования выполнен с возможностью комбинирования первого сигнала возбуждения и второго сигнала возбуждения для формирования комбинированного сигнала возбуждения для части синтезированного сигнала. Синтезатор выполнен с возможностью синтеза части синтезированного сигнала из комбинированного сигнала возбуждения и коэффициентов прогнозирования.

Дополнительные варианты осуществления второго аспекта предоставляют кодированный аудиосигнал, содержащий информацию, связанную с коэффициентами прогнозирования, информацию, связанную с детерминированной таблицей кодирования, информацию, связанную с первым параметром усиления и вторым параметром усиления, и информацию, связанную с вокализованным и невокализованным кадром сигнала.

Дополнительные варианты осуществления второго аспекта предоставляют способы для кодирования и декодирования аудиосигнала, принимаемого аудиосигнала, соответственно, и компьютерную программу.

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

Фиг. 1 показывает принципиальную блок-схему кодера для кодирования аудиосигнала согласно варианту осуществления первого аспекта;

Фиг. 2 показывает принципиальную блок-схему декодера для декодирования принимаемого входного сигнала согласно варианту осуществления первого аспекта;

Фиг. 3 показывает принципиальную блок-схему дополнительного кодера для кодирования аудиосигнала согласно варианту осуществления первого аспекта;

Фиг. 4 показывает принципиальную блок-схему кодера, содержащего отличающийся модуль вычисления параметров усиления относительно фиг. 3, согласно варианту осуществления первого аспекта;

Фиг. 5 показывает принципиальную блок-схему модуля вычисления параметров усиления, выполненного с возможностью вычисления информации первых параметров усиления и формирования сигнала с возбуждением по коду по согласно варианту осуществления второго аспекта;

Фиг. 6 показывает принципиальную блок-схему кодера для кодирования аудиосигнала, содержащего модуль вычисления параметров усиления, описанного на фиг. 5, согласно варианту осуществления второго аспекта;

Фиг. 7 показывает принципиальную блок-схему модуля вычисления параметров усиления, который содержит дополнительный формирователь, выполненный с возможностью формирования шумоподобного сигнала, относительно фиг. 5, согласно варианту осуществления второго аспекта;

Фиг. 8 показывает принципиальную блок-схему схемы невокализованного кодирования для CELP согласно варианту осуществления второго аспекта;

Фиг. 9 показывает принципиальную блок-схему параметрического невокализованного кодирования согласно варианту осуществления первого аспекта;

Фиг. 10 показывает принципиальную блок-схему декодера для декодирования кодированного аудиосигнала согласно варианту осуществления второго аспекта;

Фиг. 11a показывает принципиальную блок-схему формирователя, реализующую альтернативную структуру относительно формирователя, показанного на фиг. 2, согласно варианту осуществления первого аспекта;

Фиг. 11b показывает принципиальную блок-схему дополнительного формирователя, реализующего дополнительную альтернативу относительно формирователя,

показанного на фиг. 2, согласно варианту осуществления первого аспекта;

Фиг. 12 показывает блок-схему последовательности операций способа для кодирования аудиосигнала согласно варианту осуществления первого аспекта;

Фиг. 13 показывает блок-схему последовательности операций способа для декодирования принимаемого аудиосигнала, содержащего коэффициенты прогнозирования и параметр усиления, согласно варианту осуществления первого аспекта;

Фиг. 14 показывает блок-схему последовательности операций способа для кодирования аудиосигнала согласно варианту осуществления второго аспекта; и

Фиг. 15 показывает блок-схему последовательности операций способа для декодирования принимаемого аудиосигнала согласно варианту осуществления второго аспекта.

Идентичные или эквивалентные элементы или элементы с идентичной или эквивалентной функциональностью обозначаются в нижеприведенном описании посредством идентичных или эквивалентных ссылок с номерами даже при возникновении на различных чертежах.

В нижеприведенном описании, множество деталей изложено с тем, чтобы предоставлять более полное пояснение вариантов осуществления настоящего изобретения. Тем не менее, специалистам в данной области техники должно быть очевидным, что варианты осуществления настоящего изобретения могут быть использованы на практике без этих конкретных деталей. В других случаях, известные структуры и устройства показаны в форме блок-схемы, а не подробно, чтобы не затруднять понимание вариантов осуществления настоящего изобретения. Помимо этого, признаки различных вариантов осуществления, описанных далее, могут комбинироваться между собой, если прямо не указано иное.

Далее приводятся сведения по модификации аудиосигнала. Аудиосигнал может модифицироваться посредством усиления и/или ослабления частей аудиосигнала. Часть аудиосигнала, например, может представлять собой последовательность аудиосигнала во временной области и/или его спектра в частотной области. Относительно частотной области, спектр может модифицироваться посредством усиления или ослабления спектральных значений, размещаемых на/в частотах или частотных диапазонах. Модификация спектра аудиосигнала может содержать последовательность операций, таких как усиление и/или ослабление первой частоты или частотного диапазона и впоследствии усиление и/или ослабление второй частоты или частотного диапазона. Модификации в частотной области могут представляться как вычисление, например, умножение, деление, суммирование и т.п., спектральных значений и значений усиления и/или значений ослабления. Модификации могут выполняться последовательно, к примеру, умножение спектральных значений сначала на первое значение умножения, а затем на второе значение умножения. Умножение на второе значение умножения, а затем на первое значение умножения может обеспечивать возможность приема идентичного или почти идентичного результата. Кроме того, первое значение умножения и второе значение умножения могут сначала комбинироваться и затем применяться с точки зрения комбинированного значения умножения к спектральным значениям при приеме идентичного или сравнимого результата операции. Таким образом, этапы модификации, выполненные с возможностью формировать или модифицировать спектр аудиосигнала, описанные ниже, не ограничены описанным порядком, но также могут выполняться в измененном порядке при приеме идентичного результата и/или эффекта.

Фиг. 1 показывает принципиальную блок-схему кодера 100 для кодирования

аудиосигнала 102. Кодер 100 содержит компоновщик 110 кадров, выполненный с возможностью формировать последовательность 112 кадров на основе аудиосигнала 102. Последовательность 112 содержит множество кадров, при этом каждый кадр аудиосигнала 102 имеет длину (длительность) во временной области. Например, каждый кадр может иметь длину в 10 мс, 20 мс или 30 мс.

Кодер 100 содержит анализатор 120, выполненный с возможностью извлечения коэффициентов 122 прогнозирования (LPC – коэффициентов линейного прогнозирования) и остаточного сигнала 124 из кадра аудиосигнала. Компоновщик 110 кадров или анализатор 120 выполнен с возможностью определять представление аудиосигнала 102 в частотной области. Альтернативно, аудиосигнал 102 уже может быть представлением в частотной области.

Коэффициенты 122 прогнозирования, например, могут представлять собой коэффициенты линейного прогнозирования. Альтернативно, также нелинейное прогнозирование может применяться таким образом, что модуль 120 прогнозирования выполнен с возможностью определять коэффициенты нелинейного прогнозирования. Преимущество линейного прогнозирования заключается в уменьшенных вычислительных затратах для определения коэффициентов прогнозирования.

Кодер 100 содержит решающий модуль 130 по вокализованным/невокализованным характеристикам, выполненный с возможностью определения того, определен или нет остаточный сигнал 124 из невокализованного аудиокадра. Решающий модуль 130 выполнен с возможностью предоставления остаточного сигнала в кодер 140 вокализованных кадров, если остаточный сигнал 124 определен из вокализованного кадра сигнала, и предоставления остаточного сигнала в модуль 150 вычисления параметров усиления, если остаточный сигнал 124 определен из невокализованного аудиокадра. Для определения того, определен остаточный сигнал 122 из вокализованного или невокализованного кадра сигнала, решающий модуль 130 может использовать разные подходы, такие как автокорреляция выборок остаточного сигнала. Способ для определения того, является кадр сигнала вокализованным или невокализованным, предоставляется, например, в стандарте G.718 ITU (Международного союза по телекоммуникациям) – Т (Сектор стандартизации связи). Большая величина энергии, размещаемой на низких частотах, может указывать вокализованную часть сигнала. Альтернативно, невокализованный сигнал может приводить к большим величинам энергии на высоких частотах.

Кодер 100 содержит модуль 160 вычисления информации формант, выполненный с возможностью вычисления информации формирования речевого спектра из коэффициентов 122 прогнозирования.

Информация формирования речевого спектра может рассматривать информацию формант, например, посредством определения частот или частотных диапазонов обработанного аудиокадра, которые содержат большую величину энергии, чем окружение. Информация формирования спектра имеет возможность сегментировать спектр абсолютной величины речи на формантные (т.е. пики) и неформантные (т.е. впадины) частотные области. Области формант спектра, например, могут извлекаться посредством использования представления в форме частот спектральных иммитансов (ISF) или частот спектральных линий (LSF) коэффициентов 122 прогнозирования. Фактически, ISF или LSF представляют частоты, для которых синтезирующий фильтр с использованием коэффициентов 122 прогнозирования резонирует.

Информация 162 формирования речевого спектра и невокализованные остатки перенаправляются в модуль 150 вычисления параметров усиления, который выполнен

с возможностью вычислять параметр g_n усиления из невокализованного остаточного сигнала и информации 162 формирования спектра. Параметр g_n усиления может быть скалярным значением или их множеством, т.е. параметр усиления может содержать множество значений, связанных с усилением или ослаблением спектральных значений во множестве частотных диапазонов спектра сигнала, который должен усиливаться или ослабляться. Декодер может быть выполнен с возможностью применять параметр g_n усиления к информации принимаемого кодированного аудиосигнала таким образом, что части принимаемых кодированных аудиосигналов усиливаются или ослабляются на основе параметра усиления в ходе декодирования. Модуль 150 вычисления параметров усиления может быть выполнен с возможностью определять параметр g_n усиления посредством одного или более математических выражений или правил определения, приводящих к непрерывному значению. Операции, выполняемые в цифровой форме, например, посредством процессора, выражающие результат в переменной с ограниченным числом битов, могут приводить к квантованному усилению $\hat{g}_n$. Альтернативно, результат дополнительно может квантоваться согласно схеме квантования таким образом, что получается информация квантованного усиления. Следовательно, кодер 100 может содержать квантователь 170. Квантователь 170 может быть выполнен с возможностью квантовать определенное усиление g_n в ближайшее цифровое значение, поддерживаемое посредством цифровых операций кодера 100. Альтернативно, квантователь 170 может быть выполнен с возможностью применять функцию квантования (линейного или нелинейного) к уже оцифрованному и в силу этого квантованному коэффициенту g_n усиления. Функция нелинейного квантования может считать, например, логарифмические зависимости человеческого слуха очень чувствительными при низких уровнях звукового давления и менее чувствительными при высоких уровнях давления.

Кодер 100 дополнительно содержит модуль 180 извлечения информации, выполненный с возможностью извлечения связанной с коэффициентами прогнозирования информации 182 из коэффициентов 122 прогнозирования. Коэффициенты прогнозирования, такие как коэффициенты линейного прогнозирования, используемые для возбуждения изобретаемых таблиц кодирования, содержат низкую устойчивость к искажениям или ошибкам. Следовательно, например, известно преобразовывать коэффициенты линейного прогнозирования в межспектральные частоты (ISF) и/или извлекать пары спектральных линий (LSP) и передавать информацию, связанную с ними, вместе с кодированным аудиосигналом. Информация LSP и/или ISF содержит более высокую устойчивость к искажениям в средах передачи, например, к ошибке или ошибкам модуля вычисления. Модуль 180 извлечения информации дополнительно может содержать квантователь, выполненный с возможностью предоставлять квантованную информацию относительно LSF и/или ISP.

Альтернативно, модуль извлечения информации может быть выполнен с возможностью перенаправлять коэффициенты 122 прогнозирования. Альтернативно, кодер 100 может быть реализован без модуля 180 извлечения информации.

Альтернативно, квантователь может представлять собой функциональный блок модуля 150 вычисления параметров усиления или модуля 190 формирования потоков битов, так что модуль 190 формирования потоков битов выполнен с возможностью принимать параметр g_n усиления и извлекать квантованное усиление $\hat{g}_n$ на его основе.

Альтернативно, когда параметр g_n усиления уже квантован, кодер 100 может быть

реализован без квантователя 170.

Кодер 100 содержит модуль 190 формирования потоков битов, выполненный с возможностью принимать вокализованный сигнал, вокализованную информацию 142, связанную с вокализованным кадром кодированного аудиосигнала, соответственно, предоставленного посредством кодера 140 вокализованных кадров, принимать квантованное усиление $\hat{g}_n$ и связанную с коэффициентами прогнозирования информацию 182 и формировать выходной сигнал 192 на их основе.

Кодер 100 может представлять собой часть устройства речевого кодирования, такого как стационарный или мобильный телефон, либо устройства, содержащего микрофон для передачи аудиосигналов, такого как компьютер, планшетный РС и т.п. Выходной сигнал 192 или сигнал, извлекаемый из него, могут передаваться, например, через мобильную связь (беспроводную связь) или через проводную связь, к примеру, сетевой сигнал.

Преимущество кодера 100 состоит в том, что выходной сигнал 192 содержит информацию, извлекаемую из информации формирования спектра, преобразованной в квантованное усиление $\hat{g}_n$. Следовательно, декодирование выходного сигнала 192 может обеспечивать возможность достижения или получения дополнительной информации, которая является речевой, и в силу этого декодировать сигнал таким образом, что полученный декодированный сигнал содержит высокое качество относительно воспринимаемого уровня качества речи.

Фиг. 2 показывает принципиальную блок-схему декодера 200 для декодирования принимаемого входного сигнала 202. Принимаемый входной сигнал 202 может соответствовать, например, выходному сигналу 192, предоставленному посредством кодера 100, при этом выходной сигнал 192 может кодироваться посредством высокоуровневых кодеров, передаваться через среду, приниматься посредством приемного устройства, декодироваться на верхних уровнях, приводя к входному сигналу 202 для декодера 200.

Декодер 200 содержит модуль обратного формирования потоков битов (демультиплексор; демультиплексор) для приема входного сигнала 202. Модуль 210 обратного формирования потоков битов выполнен с возможностью предоставлять коэффициенты 122 прогнозирования, квантованное усиление $\hat{g}_n$ и вокализованную информацию 142. Для получения коэффициентов 122 прогнозирования модуль обратного формирования потоков битов может содержать модуль обратного извлечения информации, выполняющий обратную операцию относительно модуля 180 извлечения информации. Альтернативно, декодер 200 может содержать непоказанный модуль обратного извлечения информации, выполненный с возможностью выполнения обратной операции относительно модуля 180 извлечения информации. Другими словами, коэффициенты прогнозирования декодируются, т.е. восстанавливаются.

Декодер 200 содержит модуль 220 вычисления информации формант, выполненный с возможностью вычисления информации формирования речевого спектра из коэффициентов 122 прогнозирования, как описано для модуля 160 вычисления информации формант. Модуль 220 вычисления информации формант выполнен с возможностью предоставлять информацию 222 формирования речевого спектра. Альтернативно, входной сигнал 202 также может содержать информацию 222 формирования речевого спектра, при этом передача коэффициентов прогнозирования или информации, связанной с ними, такой как, например, квантованная LSF и/или ISF, вместо информации 222 формирования речевого спектра обеспечивает более низкую

скорость передачи битов входного сигнала 202.

Декодер 200 содержит генератор 240 случайного шума, выполненный с возможностью формирования шумоподобного сигнала, который может упрощенный обозначаться как шумовой сигнал. Генератор 240 случайного шума может быть выполнен с
5 возможностью воспроизводить шумовой сигнал, который получен, например, при измерении и сохранении шумового сигнала. Шумовой сигнал может измеряться и записываться, например, посредством формирования теплового шума в сопротивлении или другом электрическом компоненте и посредством сохранения записанных данных в запоминающем устройстве. Генератор 240 случайного шума выполнен с возможностью
10 предоставлять шумо(подобный) сигнал $n(n)$.

Декодер 200 содержит формирователь 250, содержащий процессор 252 формирования и регулируемый усилитель 254. Формирователь 250 выполнен с возможностью придания определенной формы спектру шумового сигнала $n(n)$. Процессор 252 формирования выполнен с возможностью приема информации формирования речевого спектра и
15 придания определенной формы спектру шумового сигнала $n(n)$, например, посредством умножения спектральных значений спектра шумового сигнала $n(n)$ и значений информации формирования спектра. Операция также может выполняться во временной области посредством свертки шумового сигнала $n(n)$ с помощью фильтра, заданного посредством информации формирования спектра. Процессор 252 формирования
20 выполнен с возможностью предоставления шумового сигнала 256 определенной формы, его спектра, соответственно, в регулируемый усилитель 254. Регулируемый усилитель 254 выполнен с возможностью приема параметра g_n усиления и усиления спектра шумового сигнала 256 определенной формы, чтобы получать усиленный шумовой сигнал 258 определенной формы. Усилитель может быть выполнен с возможностью
25 умножать спектральные значения шумового сигнала 256 определенной формы на значения параметра g_n усиления. Как указано выше, формирователь 250 может реализовываться таким образом, что регулируемый усилитель 254 выполнен с возможностью принимать шумовой сигнал $n(n)$ и предоставлять усиленный шумовой сигнал в процессор 252 формирования, выполненный с возможностью формирования
30 усиленного шумового сигнала. Альтернативно, процессор 252 формирования может быть выполнен с возможностью принимать информацию 222 формирования речевого спектра и параметр g_n усиления и применять последовательно, по одной, оба вида информации к шумовому сигналу $n(n)$ либо комбинировать оба вида информации,
35 например, посредством умножения или других вычислений и применять комбинированный параметр к шумовому сигналу $n(n)$.

Шумоподобный сигнал $n(n)$ или его усиленная версия, сформированная с информацией формирования речевого спектра, обеспечивают декодированный аудиосигнал 282, содержащий более речевое (естественное) качество звука. Это обеспечивает возможность
40 получения высококачественных аудиосигналов и/или уменьшение скоростей передачи битов на стороне кодера при поддержании или улучшении выходного сигнала 282 в декодере с уменьшенным охватом.

Декодер 200 содержит синтезатор 260, выполненный с возможностью приема коэффициентов 122 прогнозирования и усиленного шумового сигнала 258 определенной
45 формы и синтеза синтезированного сигнала 262 из усиленного шумоподобного сигнала 258 определенной формы и коэффициентов 122 прогнозирования. Синтезатор 260 может содержать фильтр и может быть выполнен с возможностью адаптации фильтра с коэффициентами прогнозирования. Синтезатор может быть выполнен с

возможностью фильтровать усиленный шумоподобный сигнал 258 определенной формы с помощью фильтра. Фильтр может реализовываться как программное обеспечение или как аппаратная структура и может содержать структуру с бесконечной импульсной характеристикой (IIR) или с конечной импульсной характеристикой (FIR).

5 Синтезированный сигнал соответствует невокализованному декодированному кадру выходного сигнала 282 декодера 200. Выходной сигнал 282 содержит последовательность кадров, которые могут преобразовываться в непрерывный аудиосигнал.

Модуль 210 обратного формирования потоков битов выполнен с возможностью разделения и предоставления сигнала 142 вокализованной информации из входного сигнала 202. Декодер 200 содержит декодер 270 вокализованных кадров, выполненный с возможностью предоставления вокализованного кадра на основе вокализованной информации 142. Декодер вокализованных кадров (процессор вокализованных кадров) выполнен с возможностью определять вокализованный сигнал 272 на основе вокализованной информации 142. Вокализованный сигнал 272 может соответствовать вокализованному аудиокадру и/или вокализованному остатку декодера 100.

Декодер 200 содержит модуль 280 комбинирования, выполненный с возможностью комбинирования невокализованного декодированного кадра 262 и вокализованного кадра 272, чтобы получать декодированный аудиосигнал 282.

Альтернативно, формирователь 250 может быть реализован без усилителя таким образом, что формирователь 250 выполнен с возможностью придания определенной формы спектру шумоподобного сигнала $p(n)$ без дополнительного усиления получаемого сигнала. Это может обеспечивать уменьшенный объем информации, передаваемый посредством входного сигнала 222, и в силу этого уменьшенную скорость передачи битов или меньшую длительность последовательности входного сигнала 202.

Альтернативно или помимо этого, декодер 200 может быть выполнен с возможностью декодировать только невокализованные кадры или обрабатывать вокализованные и невокализованные кадры посредством как спектрального формирования шумового сигнала $p(n)$, так и посредством синтеза синтезированного сигнала 262 для вокализованных и невокализованных кадров. Это может обеспечивать возможность реализации декодера 200 без декодера 270 вокализованных кадров и/или без модуля 280 комбинирования и в силу этого приводить к меньшей сложности декодера 200.

Выходной сигнал 192 и/или входной сигнал 202 содержит информацию, связанную с коэффициентами 122 прогнозирования, информацию для вокализованного кадра и невокализованного кадра, такую как флаг, указывающий то, является обработанный кадр вокализованным или невокализованным, и дополнительную информацию, связанную с вокализованным кадром сигнала, такую как кодированный вокализованный сигнал. Выходной сигнал 192 и/или входной сигнал 202 дополнительно содержит параметр усиления или параметр квантованного усиления для невокализованного кадра, так что невокализованный кадр может декодироваться на основе коэффициентов 122 прогнозирования и параметра g_n , $\hat{g}_n$ усиления, соответственно.

Фиг. 3 показывает принципиальную блок-схему кодера 300 для кодирования аудиосигнала 102. Кодер 300 содержит компоновщик 110 кадров, модуль 320 прогнозирования, выполненный с возможностью определения коэффициентов 322 линейного прогнозирования и остаточного сигнала 324 посредством применения фильтра $A(z)$ к последовательности 112 кадров, предоставленной посредством компоновщика 110 кадров. Кодер 300 содержит решающий модуль 130 и кодер 140 вокализованных кадров, чтобы получать информацию 142 вокализованных сигналов. Кодер 300 дополнительно содержит модуль 160 вычисления информации формант и

модуль 350 вычисления параметров усиления.

Модуль 350 вычисления параметров усиления выполнен с возможностью предоставления параметра g_n усиления, как описано выше. Модуль 350 вычисления параметров усиления содержит генератор 350a случайного шума для формирования шумоподобного сигнала 350b для кодирования. Модуль 350 вычисления усиления дополнительно содержит формирователь 350c, имеющий процессор 350d формирования и регулируемый усилитель 350e. Процессор 350d формирования выполнен с возможностью приема информации 222 формирования речевого спектра и шумоподобного сигнала 350b и придания определенной формы спектру шумоподобного сигнала 350b с помощью информации 162 формирования речевого спектра, как описано для формирователя 250. Регулируемый усилитель 350e выполнен с возможностью усиления шумоподобного сигнала 350f определенной формы с помощью параметра g_n (temp) усиления, который является временным параметром усиления, принимаемым из контроллера 350k. Регулируемый усилитель 350e дополнительно выполнен с возможностью предоставления усиленного шумоподобного сигнала 350g определенной формы, как описано для усиленного шумоподобного сигнала 258. Как описано для формирователя 250, порядок формирования и усиления шумоподобного сигнала может комбинироваться или изменяться относительно фиг. 3.

Модуль 350 вычисления параметров усиления содержит модуль 350h сравнения, выполненный с возможностью сравнения невокализованного остатка, предоставленного посредством решающего модуля 130, и усиленного шумоподобного сигнала 350g определенной формы. Модуль сравнения выполнен с возможностью получать показатель для сходства невокализованного остатка и усиленного шумоподобного сигнала 350g определенной формы. Например, модуль 350h сравнения может быть выполнен с возможностью определения взаимной корреляции обоих сигналов. Альтернативно или помимо этого, модуль 350h сравнения может быть выполнен с возможностью сравнения спектральных значений обоих сигналов в некоторых или всех элементах разрешения по частоте. Модуль 350h сравнения дополнительно выполнен с возможностью получать результат 350i сравнения.

Модуль 350 вычисления параметров усиления содержит контроллер 350k, выполненный с возможностью определения параметра g_n (temp) усиления на основе результата 350i сравнения. Например, когда результат 350i сравнения указывает то, что усиленный шумоподобный сигнал определенной формы содержит амплитуду или абсолютную величину, которая ниже соответствующей амплитуды или абсолютной величины невокализованного остатка, контроллер может быть выполнен с возможностью увеличивать одно или более значений параметра g_n (temp) усиления для некоторых или всех частот усиленного шумоподобного сигнала 350g. Альтернативно или помимо этого, контроллер может быть выполнен с возможностью уменьшать одно или более значений параметра g_n (temp) усиления, когда результат 350i сравнения указывает то, что усиленный шумоподобный сигнал определенной формы содержит слишком высокую абсолютную величину или амплитуду, т.е. то, что усиленный шумоподобный сигнал определенной формы является слишком громким. Генератор 350a случайного шума, формирователь 350c, модуль 350h сравнения и контроллер 350k могут быть выполнены с возможностью реализовывать оптимизацию с замкнутым контуром для определения параметра g_n (temp) усиления. Когда показатель для сходства невокализованного остатка с усиленным шумоподобным сигналом 350g определенной формы, например, выражаемого как разность между обоими сигналами, указывает то,

что сходство выше порогового значения, контроллер 350k выполнен с возможностью предоставлять определенный параметр g_n усиления. Квантователь 370 выполнен с возможностью квантовать параметр g_n усиления, чтобы получать параметр $\hat{g}_n$ квантованного усиления.

Генератор 350a случайного шума может быть выполнен с возможностью доставлять гауссов шум. Генератор 350a случайного шума может быть выполнен с возможностью выполнения (вызова) генератора случайных чисел с числом n равномерных распределений между нижним пределом (минимальным значением), к примеру, -1 и верхним пределом (максимальным значением), к примеру, +1. Например, генератор случайного шума 350 выполнен с возможностью вызова три раза генератора случайных чисел. Поскольку реализованные в цифровой форме генераторы случайного шума могут выводить псевдослучайные значения, добавление либо наложение нескольких или множества псевдослучайных функций может обеспечивать возможность получения достаточно случайной распределенной функции. Эта процедура подчиняется центральной предельной теореме. Генератор 350a случайного шума может быть выполнен с возможностью вызывать генератор случайных чисел, по меньшей мере, два, три или более раз, как указано посредством следующего псевдокода:

```
for(i=0;i<Ls;i++){
  n[i]=uniform_random;
  n[i]+=uniform_random;
  n[i]+=uniform_random;
}
```

Альтернативно, генератор 350a случайного шума может формировать шумоподобный сигнал из запоминающего устройства, как описано для генератора 240 случайного шума. Альтернативно, генератор 350a случайного шума может содержать, например, электрическое сопротивление или другое средство для формирования шумового сигнала посредством выполнения кода или посредством измерения физических эффектов, таких как тепловой шум.

Процессор 350b формирования может быть выполнен с возможностью добавлять формантную структуру и наклон в шумоподобные сигналы 350b посредством фильтрации шумоподобного сигнала 350b с помощью $fe(n)$, как указано выше. Наклон может добавляться посредством фильтрации сигнала с помощью фильтра $t(n)$, содержащего передаточную функцию, на основе следующего:

$$Ft(z) = 1 - \beta z^{-1},$$

где коэффициент β может быть выведен из вокализации предыдущего субкадра:

$$voicing = \frac{energy(contibution\ of\ AC) - energy(contibution\ of\ IC)}{energy(sum\ of\ contributions)}$$

где AC является сокращением для адаптивной таблицы кодирования, а IC является сокращением для изобретаемой таблицы кодирования.

$$\beta = 0.25 \cdot (1 + voicing),$$

Параметр g_n усиления, параметр $\hat{g}_n$ квантованного усиления, соответственно, обеспечивают предоставление дополнительной информации, что позволяет уменьшать ошибку или рассогласование между кодированным сигналом и соответствующим декодированным сигналом, декодированным в декодере, к примеру, в декодере 200.

Относительно правила определения:

$$Ffe(z) = \frac{A(z/w1)}{A(z/w2)},$$

параметр w1 может содержать положительное ненулевое значение самое большее в 1,0, предпочтительно, самое меньшее в 0,7 и самое большее в 0,8, и более
 5 предпочтительно содержать значение в 0,75. Параметр w2 может содержать положительное ненулевое скалярное значение самое большее в 1,0, предпочтительно, самое меньшее в 0,8 и самое большее в 0,93, и более предпочтительно содержать значение в 0,9. Параметр w2 предпочтительно превышает w1.

Фиг. 4 показывает принципиальную блок-схему кодера 400. Кодер 400 выполнен с
 10 возможностью предоставлять информацию 142 вокализованных сигналов, как описано для кодеров 100 и 300. По сравнению с кодером 300, кодер 400 содержит отличающийся модуль 350' вычисления параметров усиления. Модуль 350h' сравнения выполнен с возможностью сравнивать аудиокадр 112 и синтезированный сигнал 350l', чтобы
 15 получать результат 350i' сравнения. Модуль 350' вычисления параметров усиления содержит синтезатор 350m', выполненный с возможностью синтеза синтезированного сигнала 350l' на основе усиленного шумоподобного сигнала 350g определенной формы и коэффициентов 122 прогнозирования.

По существу, модуль 350' вычисления параметров усиления реализует, по меньшей мере, частично декодер посредством синтеза синтезированного сигнала 350l'.
 20 По сравнению с кодером 300, содержащим модуль 350h сравнения, выполненный с возможностью сравнения невокализованного остатка и усиленного шумоподобного сигнала определенной формы, кодер 400 содержит модуль 350h' сравнения, который выполнен с возможностью сравнивать (вероятно, полный) аудиокадр и синтезированный сигнал. Это позволяет обеспечивать более высокую точность, поскольку кадры сигнала,
 25 а не только их параметры, сравниваются между собой. Более высокая точность может требовать увеличенных вычислительных затрат, поскольку аудиокадр 122 и синтезированный сигнал 350l' могут содержать более высокую сложность по сравнению с остаточным сигналом и с усиленной шумоподобной информацией определенной формы, так что сравнение обоих сигналов также является более сложным. Помимо
 30 этого, должен вычисляться синтез, что требует вычислительных затрат посредством синтезатора 350m'.

Модуль 350' вычисления параметров усиления содержит запоминающее устройство 350n', выполненное с возможностью записи информации кодирования, содержащей
 35 параметр g_n усиления при кодировании или его квантованную версию $\hat{g}_n$. Это дает возможность контроллеру 350k получать сохраненное значение усиления при обработке последующего аудиокадра. Например, контроллер может быть выполнен с возможностью определять первое (набор) значение(я), т.е. первый экземпляр коэффициента $g_n(\text{temp})$ усиления на основе или равного значению g_n для предыдущего
 40 аудиокадра.

Фиг. 5 показывает принципиальную блок-схему модуля 550 вычисления параметров усиления, выполненного с возможностью вычисления информации g_n первых параметров
 усиления согласно второму аспекту. Модуль 550 вычисления параметров усиления содержит генератор 550a сигналов, выполненный с возможностью формирования
 45 сигнала $s(n)$ возбуждения. Генератор 550a сигналов содержит детерминированную таблицу кодирования и индекс в таблице кодирования, чтобы формировать сигнал $s(n)$. Иными словами, входная информация, такая как коэффициенты 122 прогнозирования, приводит к детерминированному сигналу $s(n)$ возбуждения. Генератор

550a сигналов может быть выполнен с возможностью формировать сигнал $s(n)$ возбуждения согласно изобретаемой таблице кодирования схемы CELP-кодирования. Таблица кодирования может определяться или обучаться согласно измеренным речевым данным на предыдущих этапах калибровки. Модуль вычисления параметров усиления содержит формирователь 550b, выполненный с возможностью придания определенной формы спектру кодового сигнала $s(n)$ на основе информации 550c формирования речевого спектра для кодового сигнала $s(n)$. Информация 550c формирования речевого спектра может получаться из контроллера 160 информации формант. Формирователь 550b содержит процессор 550d формирования, выполненный с возможностью приема информации 550c формирования для формирования кодового сигнала. Формирователь 550b дополнительно содержит регулируемый усилитель 550e, выполненный с возможностью усиления кодового сигнала $s(n)$ определенной формы, чтобы получать усиленный кодовый сигнал 550f определенной формы. Таким образом, параметр кодового усиления выполнен с возможностью задания кодового сигнала $s(n)$, который связан с детерминированной таблицей кодирования.

Модуль 550 вычисления параметров усиления содержит генератор 350a шума, выполненный с возможностью предоставления шумо(подобного) сигнала $p(n)$, и усилитель 550g, выполненный с возможностью усиления шумового сигнала $p(n)$ на основе параметра g_n усиления шума, чтобы получать усиленный шумовой сигнал 550h.

Модуль вычисления параметров усиления содержит модуль 550i комбинирования, выполненный с возможностью комбинирования усиленного кодового сигнала 550f определенной формы и усиленного шумового сигнала 550h, чтобы получать комбинированный сигнал 550k возбуждения. Модуль 550i комбинирования, например, может быть выполнен с возможностью спектрального суммирования или умножения спектральных значений усиленного кодового сигнала определенной формы и усиленного шумового сигнала 550f и 550h. Альтернативно, модуль 550i комбинирования может быть выполнен с возможностью свертывать оба сигнала 550f и 550h.

Как описано выше для формирователя 350c, формирователь 550b может реализовываться таким образом, что кодовый сигнал $s(n)$ сначала усиливается посредством регулируемого усилителя 550e и далее формируется посредством процессора 550d формирования. Альтернативно, информация 550c формирования для кодового сигнала $s(n)$ может комбинироваться с информацией g_c параметров кодового усиления таким образом, что комбинированная информация применяется к кодовому сигналу $s(n)$.

Модуль 550 вычисления параметров усиления содержит модуль 550l сравнения, выполненный с возможностью сравнения комбинированного сигнала 550k возбуждения и невокализованного остаточного сигнала, полученного для решающего модуля 130 по вокализованным/невокализованным характеристикам. Модуль 550l сравнения может представлять собой модуль 550h сравнения и может быть выполнен с возможностью предоставления результата сравнения, т.е. показателя 550m для сходства комбинированного сигнала 550k возбуждения и невокализованного остаточного сигнала. Модуль вычисления кодового усиления содержит контроллер 550n, выполненный с возможностью управления информацией g_c параметров кодового усиления и информацией g_n параметров усиления шума. Параметр g_c кодового усиления и информация g_n параметров усиления шума могут содержать несколько или множество скалярных или мнимых значений, которые могут быть связаны с частотным диапазоном шумового сигнала $p(n)$ или сигнала, извлекаемого из него, либо со спектром кодового

сигнала $s(n)$ или сигнала, извлекаемого из него.

Альтернативно, модуль 550 вычисления параметров усиления может реализовываться без процессора 550d формирования. Альтернативно, процессор 550d формирования может быть выполнен с возможностью формировать шумовой сигнал $n(n)$ и предоставлять шумовой сигнал определенной формы в регулируемый усилитель 550g.

Таким образом, посредством управления обеими видами информации g_c и g_n параметров усиления, сходство комбинированного сигнала 550k возбуждения относительно невокализованного остатка может повышаться, так что декодер, принимающий информацию в информацию g_c параметров кодового усиления и информацию g_n параметров усиления шума, может воспроизводить аудиосигнал, который содержит хорошее качество звука. Контроллер 550n выполнен с возможностью предоставлять выходной сигнал 550o, содержащий информацию, связанную с информацией g_c параметров кодового усиления и информацией g_n параметров усиления шума. Например, сигнал 550o может содержать оба вида информации g_n и g_c параметров усиления в качестве скалярных или квантованных значений либо в качестве значений, извлеченных из них, например, кодированных значений.

Фиг. 6 показывает принципиальную блок-схему кодера 600 для кодирования аудиосигнала 102, содержащего модуль 550 вычисления параметров усиления, описанный на фиг. 5. Кодер 600 может получаться, например, посредством модификации кодера 100 или 300. Кодер 600 содержит первый квантователь 170-1 и второй квантователь 170-2. Первый квантователь 170-1 выполнен с возможностью квантования информации g_c параметров усиления для получения информации $\hat{g}_c$ квантованных параметров усиления. Второй квантователь 170-2 выполнен с возможностью квантования информации g_n параметров усиления шума для получения информации $\hat{g}_n$ квантованных параметров усиления шума. Модуль 690 формирования потоков битов выполнен с возможностью формирования выходного сигнала 692, содержащего информацию 142 вокализованных сигналов, связанную с LPC информацией 122 и информацией $\hat{g}_c$ и $\hat{g}_n$ квантованных параметров усиления. По сравнению с выходным сигналом 192, выходной сигнал 692 расширяется или обновляется посредством информации $\hat{g}_c$ квантованных параметров усиления. Альтернативно, квантователь 170-1 и/или 170-2 может представлять собой часть модуля 550 вычисления параметров усиления. Дополнительно, один из квантователей 170-1 и/или 170-2 может быть выполнен с возможностью получать оба параметра $\hat{g}_c$ и $\hat{g}_n$ квантованного усиления.

Альтернативно, кодер 600 может быть выполнен с возможностью содержать один квантователь, выполненный с возможностью квантования информации g_c параметров кодового усиления и параметра g_n усиления шума для получения информации $\hat{g}_c$ и $\hat{g}_n$ квантованных параметров. Оба вида информации параметров усиления могут квантоваться, например, последовательно.

Модуль 160 вычисления информации формант выполнен с возможностью вычислять информацию 550с формирования речевого спектра из коэффициентов 122 прогнозирования.

Фиг. 7 показывает принципиальную блок-схему модуля 550' вычисления параметров усиления, который модифицируется относительно модуля 550 вычисления параметров усиления. Модуль 550' вычисления параметров усиления содержит формирователь 350,

описанный на фиг. 3, вместо усилителя 550g. Формирователь 350 выполнен с возможностью предоставлять усиленный шумовой сигнал 350g определенной формы. Модуль 550i комбинирования выполнен с возможностью комбинировать усиленный кодовый сигнал 550f определенной формы и усиленный шумовой сигнал 350g определенной формы, чтобы предоставлять комбинированный сигнал 550k' возбуждения. Модуль 160 вычисления информации формант выполнен с возможностью предоставлять оба вида информации 162 и 550с речевых формант. Информация 550с и 162 речевых формант может быть одинаковой. Альтернативно, оба вида информации 550с и 162 могут отличаться друг от друга. Это обеспечивает возможность отдельного моделирования, т.е. формирования сигнала $s(n)$ и $p(n)$ с формированием на основе кода.

Контроллер 550n может быть выполнен с возможностью определения информации g_c и g_n параметров усиления для каждого субкадра обработанного аудиокадра.

Контроллер может быть выполнен с возможностью определять, т.е. вычислять, информацию g_c и g_n параметров усиления на основе подробностей, изложенных ниже.

Во-первых, средняя энергия субкадра может вычисляться для исходного сигнала остатка кратковременного прогнозирования, доступного во время LPC-анализа, т.е. для невокализованного остаточного сигнала. Энергия усредняется по четырем субкадрам текущего кадра в логарифмической области следующим образом:

$$nrg = \frac{10}{4} * \sum_{l=0}^3 \log_{10} \left(\sum_{n=0}^{Lsf-1} \frac{res^2(l \cdot Lsf + n)}{Lsf} \right),$$

где Lsf является размером субкадра в выборках. В этом случае, кадр разделен на 4 субкадра. Усредненная энергия затем может кодироваться в определенном числе битов, например, в трех, четырех или пяти, посредством использования предварительно обученной стохастической таблицы кодирования. Стохастическая таблица кодирования может содержать число записей (размер) согласно числу различных значений, которые могут быть представлены посредством числа битов, например, размер 8 для числа 3 битов, размер 16 для числа 4 битов или число 32 для числа 5 битов. Квантованное усиление $\hat{nrg}$ может определяться из выбранного кодового слова таблицы кодирования. Для каждого субкадра вычисляются два вида информации g_c и g_n усиления. Усиление g_c кода может вычисляться, например, на основе следующего:

$$g_c = \frac{\sum_{n=0}^{Lsf-1} xw(n) \cdot cw(n)}{\sum_{n=0}^{Lsf-1} cw(n) \cdot cw(n)}$$

где $cw(n)$, например, является фиксированным новшеством, выбранным из фиксированной таблицы кодирования, состоящей из генератора 550a сигналов, фильтруемого посредством перцепционного взвешивающего фильтра. Выражение $xw(n)$ соответствует традиционному целевому перцепционному возбуждению, вычисленному в CELP-кодерах. Информация g_c кодового усиления затем может быть нормализована для получения нормализованного усиления g_{nc} на основе следующего:

$$g_{nc} = g_c \cdot \frac{\sum_{n=0}^{Lsf-1} \sqrt{c(n) \cdot c(n)}}{Lsf * 10^{\hat{nrg}/20}}$$

Нормализованное усиление g_{nc} может квантоваться, например, посредством квантователя 170-1. Квантование может выполняться согласно линейной или логарифмической шкале. Логарифмическая шкала может содержать шкалу размера в 4, 5 или более битов. Например, логарифмическая шкала содержит размер в 5 битов.

Квантование может выполняться на основе следующего:

$$Index_{nc} = \lfloor 20 * \log_{10}((g_{nc} + 20)/1.25) + 0.5 \rfloor$$

где $Index_{nc}$ может быть ограничен между 0 и 31, если логарифмическая шкала содержит 5 битов. $Index_{nc}$ может быть информацией квантованных параметров усиления.

Квантованное усиление кода $\hat{g}_c$ затем может выражаться на основе следующего:

$$\hat{g}_c = 10^{10(index_{nc} - 20)/20} \cdot \frac{Lsf * 10^{\hat{r}g/20}}{\sum_{n=0}^{Lsf-1} \sqrt{c(n) \cdot c(n)}}$$

Усиление кода может вычисляться для того, чтобы минимизировать среднеквадратичную ошибку или среднеквадратическую ошибку (MSE):

$$\frac{1}{Lsf} \sum_{n=0}^{Lsf-1} (xw(n) - g_c \cdot cw(n))^2,$$

где Lsf соответствует частотам спектральных линий, определенным из коэффициентов 122 прогнозирования.

Информация параметров усиления шума может определяться с точки зрения несовпадения энергии посредством минимизации ошибки на основе следующего:

$$\frac{1}{Lsf} \left| \sum_{n=0}^{Lsf-1} k \cdot xw^2(n) - \sum_{n=0}^{Lsf-1} (\hat{g}_c \cdot cw(n) + g_n nw(n))^2 \right|$$

Переменная k является коэффициентом ослабления, который может варьироваться в зависимости или на основе коэффициентов прогнозирования, при этом коэффициенты прогнозирования могут обеспечивать определение того, содержит или нет речь низкую часть фонового шума или даже вообще не содержит фоновый шум (чистая речь).

Альтернативно, сигнал также может определяться в качестве зашумленной речи, например, когда аудиосигнал или его кадр содержит изменения между

невокализованными и невокализованными кадрами. Переменная k может задаваться равной значению самое меньшее в 0,85, самое меньшее в 0,95 или даже значению в 1

для чистой речи, когда высокая динамика энергии является перцепционно важной.

Переменная k может задаваться равной значению самое меньшее в 0,6 и самое большее в 0,9, предпочтительно значению самое меньшее в 0,7 и самое большее в 0,85, и более предпочтительно, значению в 0,8 для зашумленной речи, когда шумовое возбуждение задается более умеренным для недопущения флуктуации в выходной энергии между невокализованными и невокализованными кадрами. Ошибка (несовпадение энергии)

может вычисляться для каждого из этих возможных вариантов $\hat{g}_c$ квантованного

усиления. Кадр, разделенный на четыре субкадра, может приводить к четырем

возможным вариантам $\hat{g}_c$ квантованного усиления. Один возможный вариант, который

минимизирует ошибку, может выводиться посредством контроллера. Квантованное усиление шума (информация параметров усиления шума) может вычисляться на основе следующего:

$$\hat{g}_n = (index_n \cdot 0.25 + 0.25) \cdot \hat{g}_c \cdot \frac{\sum_{n=0}^{Lsf-1} \sqrt{c(n) \cdot c(n)}}{\sum_{n=0}^{Lsf-1} \sqrt{n(n) \cdot n(n)}},$$

где $Index_n$ ограничен между 0 и 3 согласно четырем возможным вариантам.

Результирующий комбинированный сигнал возбуждения, к примеру, сигнал 550k или 550k' возбуждения может получаться на основе следующего:

$$e(n) = \hat{g}_c \cdot c(n) + \hat{g}_n \cdot n(n),$$

где $e(n)$ является комбинированным сигналом 550k или 550k' возбуждения.

Кодер 600 или модифицированный кодер 600, содержащий модуль 550 или 550' вычисления параметров усиления, может обеспечивать возможность невокализованного кодирования на основе схемы CELP-кодирования. Схема CELP-кодирования может модифицироваться на основе следующих примерных подробностей для обработки невокализованных кадров.

Параметры LTP не передаются, поскольку фактически отсутствует периодичность в невокализованных кадрах, и результирующее усиление при кодировании является очень низким. Адаптивное возбуждение задается равным нулю.

Сэкономленные биты сообщаются в фиксированную таблицу кодирования. Большее число импульсов может кодироваться для идентичной скорости передачи битов, и в таком случае может повышаться качество.

На низких скоростях, т.е. для скоростей между 6 и 12 Кбит/с, импульсное кодирование не является достаточным для моделирования надлежащим образом целевого шумоподобного возбуждения невокализованного кадра. Гауссова таблица кодирования добавляется в фиксированную таблицу кодирования для компоновки конечного возбуждения.

Фиг. 8 показывает принципиальную блок-схему схемы невокализованного кодирования для CELP согласно второму аспекту. Модифицированный контроллер 810 содержит функции как модуля 550l сравнения, так и контроллера 550n. Контроллер 810 выполнен с возможностью определения информации g_c параметров кодового усиления и информации g_n параметров усиления шума на основе анализа через синтез, т.е. посредством сравнения синтезированного сигнала с входным сигналом, указываемым в качестве $s(n)$, который, например, представляет собой невокализованный остаток. Контроллер 810 содержит фильтр 820 по методу анализа через синтез, выполненный с возможностью формирования возбуждения для генератора 550a сигналов (изобретаемого возбуждения) и предоставления информации g_c и g_n параметров усиления.

Блок 810 по методу анализа через синтез выполнен с возможностью сравнивать комбинированный сигнал 550k' возбуждения посредством сигнала, внутренне синтезированного посредством адаптации фильтра в соответствии с предоставленными параметрами и информацией.

Контроллер 810 содержит блок анализа, выполненный с возможностью получения коэффициентов прогнозирования, как описано для анализатора 320, чтобы получать коэффициенты 122 прогнозирования. Контроллер дополнительно содержит синтезирующий фильтр 840 для фильтрации комбинированного сигнала 550k возбуждения с помощью синтезирующего фильтра 840, при этом синтезирующий фильтр 840 адаптирован посредством коэффициентов 122 фильтрации. Дополнительный модуль сравнения может быть выполнен с возможностью сравнивать входной сигнал $s(n)$ и синтезированный сигнал $\hat{s}(n)$, например, декодированный (восстановленный) аудиосигнал. Дополнительно, приспособляется запоминающее устройство 350n, при этом контроллер 810 выполнен с возможностью сохранять прогнозируемый сигнал и/или прогнозные коэффициенты в запоминающем устройстве. Генератор 850 сигналов выполнен с возможностью предоставлять сигнал адаптивного возбуждения на основе сохраненных прогнозирований в запоминающем устройстве 350n, обеспечивая возможность улучшения адаптивного возбуждения на основе первого комбинированного сигнала возбуждения.

Фиг. 9 показывает принципиальную блок-схему параметрического невокализованного кодирования согласно первому аспекту. Усиленный шумовой сигнал определенной формы может представлять собой входной сигнал синтезирующего фильтра 910, который адаптирован посредством определенных коэффициентов 122 фильтрации (коэффициентов прогнозирования). Вывод синтезированного сигнала 912 посредством синтезирующего фильтра может сравниваться с входным сигналом $s(n)$, который, например, может представлять собой аудиосигнал. Синтезированный сигнал 912 содержит ошибку относительно входного сигнала $s(n)$. Посредством модификации параметра g_n усиления шума посредством блока 920 анализа, который может соответствовать модулю 150 или 350 вычисления параметров усиления, может уменьшаться или минимизироваться ошибка. Посредством сохранения усиленного шумового сигнала 350f определенной формы в запоминающем устройстве 350n может выполняться обновление адаптивной таблицы кодирования, так что обработка вокализованных аудиокадров также может совершенствоваться на основе улучшенного кодирования невокализованного аудиокадра.

Фиг. 10 показывает принципиальную блок-схему декодера 1000 для декодирования кодированного аудиосигнала, например, кодированного аудиосигнала 692. Декодер 1000 содержит генератор 1010 сигналов и генератор 1020 шума, выполненный с возможностью формирования шумоподобного сигнала 1022. Принимаемый сигнал 1002 содержит связанную с LPC информацию, при этом модуль 1040 обратного формирования потоков битов выполнен с возможностью предоставлять коэффициенты 122 прогнозирования на основе связанной с коэффициентами прогнозирования информации. Например, декодер 1040 выполнен с возможностью извлекать коэффициенты 122 прогнозирования. Генератор 1010 сигналов выполнен с возможностью формировать сигнал 1012 возбуждения с возбуждением по коду, как описано для генератора 558 сигналов. Модуль 1050 комбинирования декодера 1000 выполнен с возможностью комбинирования сигнала 1012 с возбуждением по коду и шумоподобного сигнала 1022, как описано для модуля 550 комбинирования, чтобы получать комбинированный сигнал 1052 возбуждения. Декодер 1000 содержит синтезатор 1060, имеющий фильтр для адаптации с коэффициентами 122 прогнозирования, при этом синтезатор выполнен с возможностью фильтрации комбинированного сигнала 1052 возбуждения с помощью адаптированного фильтра, чтобы получать невокализованный декодированный кадр 1062. Декодер 1000 также содержит модуль 284 комбинирования, комбинирующий невокализованный декодированный кадр и вокализованный кадр 272, чтобы получать последовательность 282 аудиосигналов. По сравнению с декодером 200, декодер 1000 содержит генератор вторых сигналов, выполненный с возможностью предоставлять сигнал 1012 возбуждения с возбуждением по коду. Шумоподобный сигнал 1022 возбуждения, например, может представлять собой шумоподобный сигнал $p(n)$, проиллюстрированный на фиг. 2.

Последовательность 282 аудиосигналов может содержать хорошее качество и сильное сходство относительно кодированного входного сигнала.

Дополнительные варианты осуществления предоставляют декодеры, улучшающие декодер 1000 посредством формирования и/или усиления сигнала 1012 возбуждения с формированием на основе кода (с возбуждением по коду) и/или шумоподобного сигнала 1022. Таким образом, декодер 1000 может содержать процессор формирования и/или регулируемый усилитель, размещаемый между генератором 1010 сигналов и модулем 1050 комбинирования, между генератором 1020 шума и модулем 1050 комбинирования, соответственно. Входной сигнал 1002 может содержать информацию, связанную с

информацией g_c параметров кодового усиления и/или информацией параметров усиления шума, при этом декодер может быть выполнен с возможностью адаптировать усилитель для усиления сигнала 1012 возбуждения с формированием на основе кода или его версии определенной формы посредством использования информации g_c параметров кодового усиления. Альтернативно или помимо этого, декодер 1000 может быть выполнен с

возможностью адаптировать, т.е. управлять усилителем для усиления шумоподобного сигнала 1022 или его версии определенной формы с помощью усилителя посредством использования информации параметров усиления шума.

Альтернативно, декодер 1000 может содержать формирователь 1070, выполненный с возможностью формирования сигнала 1012 возбуждения с возбуждением по коду, и/или формирователь 1080, выполненный с возможностью формирования шумоподобного сигнала 1022, как указано посредством пунктирных линий. Формирователи 1070 и/или 1080 могут принимать параметры g_c и/или g_n усиления и/или информацию формирования речевого спектра. Формирователи 1070 и/или 1080 могут формироваться так, как описано для вышеописанных формирователей 250, 350с и/или 550b.

Декодер 1000 может содержать модуль 1090 вычисления формантной информации, чтобы предоставлять информацию 1092 формирования речевого спектра для формирователей 1070 и/или 1080, как описано для модуля 160 вычисления информации формант. Модуль 1090 вычисления информации формант может быть выполнен с возможностью предоставлять различную информацию (1092a; 1092b) формирования речевого спектра в формирователи 1070 и/или 1080.

Фиг. 11a показывает принципиальную блок-схему формирователя 250', реализующую альтернативную структуру по сравнению с формирователем 250. Формирователь 250' содержит модуль 257 комбинирования для комбинирования информации 222 формирования и связанного с шумом параметра g_n усиления, чтобы получать комбинированную информацию 259. Модифицированный процессор 252' формирования выполнен с возможностью формировать шумоподобный сигнал $p(n)$ посредством использования комбинированной информации 259, чтобы получать усиленный шумоподобный сигнал 258 определенной формы. Поскольку как информация 222 формирования, так и параметр g_n усиления могут быть интерпретированы в качестве коэффициентов умножения, оба коэффициента умножения могут умножаться посредством использования модуля 257 комбинирования и затем применяться в комбинированной форме к шумоподобному сигналу $p(n)$.

Фиг. 11b показывает принципиальную блок-схему формирователя 250'', реализующую дополнительную альтернативу по сравнению с формирователем 250. По сравнению с формирователем 250, регулируемый усилитель 254 выполнен с возможностью сначала формировать усиленный шумоподобный сигнал посредством усиления шумоподобного сигнала $p(n)$ с использованием параметра g_n усиления. Процессор 252 формирования выполнен с возможностью формировать усиленный сигнал с использованием информации 222 формирования, чтобы получать усиленный сигнал 258 определенной формы.

Хотя фиг. 11a и 11b связаны с формирователем 250, иллюстрирующим альтернативные реализации, вышеприведенные описания также применяются к формирователям 350с, 550b, 1070 и/или 1080.

Фиг. 12 показывает блок-схему последовательности операций способа 1200 для кодирования аудиосигнала согласно первому аспекту. Способ 1210 содержит извлечение коэффициентов прогнозирования и остаточного сигнала из кадра аудиосигнала. Способ

1200 содержит этап 1230, на котором параметр усиления вычисляется из невокализованного остаточного сигнала и информации формирования спектра, и этап 1240, на котором выходной сигнал формируется на основе информации, связанной с вокализованным кадром сигнала, параметром усиления или параметром квантованного

5 усиления и коэффициентами прогнозирования.

Фиг. 13 показывает блок-схему последовательности операций способа 1300 для декодирования принимаемого аудиосигнала, содержащего коэффициенты прогнозирования и параметр усиления, согласно первому аспекту. Способ 1300 содержит этап 1310, на котором информация формирования речевого спектра вычисляется из

10 коэффициентов прогнозирования. На этапе 1320, формируется шумоподобный сигнал для декодирования. На этапе 1330, спектр шумоподобного сигнала для декодирования или его усиленное представление формируется с использованием информации формирования спектра, чтобы получать шумоподобный сигнал для декодирования определенной формы. На этапе 1340 способа 1300, синтезированный сигнал

15 синтезируется из усиленного шумоподобного сигнала для кодирования определенной формы и коэффициентов прогнозирования.

Фиг. 14 показывает блок-схему последовательности операций способа 1400 для кодирования аудиосигнала согласно второму аспекту. Способ 1400 содержит этап 1410, на котором коэффициенты прогнозирования и остаточный сигнал извлекаются из

20 невокализованного кадра аудиосигнала. На этапе 1420 способа 1400, информация первых параметров усиления для задания первого сигнала возбуждения, связанного с детерминированной таблицей кодирования, и информация вторых параметров усиления для задания второго сигнала возбуждения, связанного с шумоподобным сигналом, вычисляется для невокализованного кадра.

На этапе 1430 способа 1400, выходной сигнал формируется на основе информации, связанной с вокализованным кадром сигнала, информации первых параметров усиления и информации вторых параметров усиления.

Фиг. 15 показывает блок-схему последовательности операций способа 1500 для декодирования принимаемого аудиосигнала согласно второму аспекту. Принимаемый

30 аудиосигнал содержит информацию, связанную с коэффициентами прогнозирования. Способ 1500 содержит этап 1510, на котором первый сигнал возбуждения формируется из детерминированной таблицы кодирования для части синтезированного сигнала. На этапе 1520 способа 1500, второй сигнал возбуждения формируется из шумоподобного сигнала для части синтезированного сигнала. На этапе 1530 способа 1000, первый

35 сигнал возбуждения и второй сигнал возбуждения комбинируются для формирования комбинированного сигнала возбуждения для части синтезированного сигнала. На этапе 1540 способа 1500, часть синтезированного сигнала синтезируется из комбинированного сигнала возбуждения и коэффициентов прогнозирования.

Другими словами, аспекты настоящего изобретения предлагают новый способ

40 кодирования невокализованных кадров посредством формирования случайно сформированного гауссова шума и придания определенной формы его спектру посредством добавления в него формантной структуры и спектрального наклона. Формирование спектра выполняется в области возбуждения перед возбуждением синтезирующего фильтра. Как следствие, возбуждение определенной формы обновляется

45 в запоминающем устройстве долговременного прогнозирования для формирования последующих адаптивных таблиц кодирования.

Последующие кадры, которые не являются невокализованными, также должны извлекать выгоду из формирования спектра. В отличие от улучшения формант при

постфильтрации, предложенное формирование шума выполняется на сторонах кодера и декодера.

Такое возбуждение может использоваться непосредственно в схеме параметрического кодирования для фокусирования на очень низких скоростях передачи битов. Тем не менее, также предлагается ассоциировать такое возбуждение в сочетании с традиционной изобретаемой таблицей кодирования в схеме CELP-кодирования.

Для обоих способов предлагается новое кодирование усиления, в частности, эффективное как для чистой речи, так и для речи с фоновым шумом. Предлагаются некоторые механизмы, чтобы максимально приближаться к исходной энергии при одновременном недопущении слишком резких переходов с неневокализованными кадрами, а также недопущении нежелательной нестабильности вследствие квантования по усилению.

Первый аспект фокусируется на невокализованном кодировании со скоростью 2,8 и 4 килобита в секунду (Кбит/с). Сначала обнаруживаются невокализованные кадры. Это может выполняться посредством обычной классификации речи, которая осуществляется в стандарте многорежимного широкополосного кодирования с переменной скоростью (VMR-WB), как известно из [3].

Предусмотрено два основных преимущества при выполнении формирования спектра на этой стадии. Во-первых, формирование спектра учитывает вычисление усиления возбуждения. Поскольку вычисление усиления представляет собой единственный неслепой модуль во время формирования возбуждения, его наличие в конце цепочки после формирования является большим преимуществом. Во-вторых, оно обеспечивает сохранение усовершенствованного возбуждения в запоминающем устройстве LTP. В таком случае улучшение также подходит для последующих неневокализованных кадров.

Хотя квантователи 170, 170-1 и 170-2 описаны как выполненные с возможностью получения квантованных параметров $\hat{g}_c$ и $\hat{g}_n$, квантованные параметры могут предоставляться в качестве информации, связанной с ними, например, индекса или идентификатора записи базы данных, причем запись содержит параметры $\hat{g}_c$ и $\hat{g}_n$ квантованного усиления.

Хотя некоторые аспекты описаны в контексте устройства, очевидно, что эти аспекты также представляют описание соответствующего способа, при этом блок или устройство соответствует этапу способа либо признаку этапа способа. Аналогично, аспекты, описанные в контексте этапа способа, также представляют описание соответствующего блока или элемента, или признака соответствующего устройства.

Изобретаемый кодированный аудиосигнал может быть сохранен на цифровом носителе хранения данных или может быть передан по среде передачи, такой как беспроводная среда передачи или проводная среда передачи, к примеру, Интернет.

В зависимости от определенных требований к реализации, варианты осуществления изобретения могут быть реализованы в аппаратных средствах или в программном обеспечении. Реализация может выполняться с использованием цифрового носителя хранения данных, например, гибкого диска, DVD, CD, ROM, PROM, EPROM, EEPROM или флэш-памяти, имеющего сохраненные электронночитаемые управляющие сигналы, которые взаимодействуют (или допускают взаимодействие) с программируемой компьютерной системой таким образом, что осуществляется соответствующий способ.

Некоторые варианты осуществления согласно изобретению содержат носитель данных, имеющий электронночитаемые управляющие сигналы, которые допускают взаимодействие с программируемой компьютерной системой таким образом, что

осуществляется один из способов, описанных в данном документе.

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

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

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

Следовательно, дополнительный вариант осуществления изобретаемых способов 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.

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, выполненное 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.

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 посредством объема нижеприведенной формулы изобретения, а не посредством конкретных подробностей, представленных посредством описания и пояснения вариантов осуществления в данном документе.

Документы

[1] Recommendation ITU-T G.718: "Frame error robust narrow-band and wideband embedded variable bit-rate coding of speech and audio from 8-32 kbit/s".

[2] Патент (США) номер US 5444816, "Dynamic codebook for efficient speech coding

based on algebraic codes".

[3] Jelinek, M.; Salami, R., "Wideband Speech Coding Advances in VMR-WB Standard", Audio, Speech and Language Processing, IEEE Transactions on, издание 15, номер 4, стр. 1167, 1179, май 2007 года.

5

(57) Формула изобретения

1. Кодер (100; 200; 300) для кодирования аудиосигнала (102), причем кодер содержит:
- анализатор (120; 320), выполненный с возможностью извлечения коэффициентов (122; 322) прогнозирования и остаточного сигнала (124; 324) из кадра аудиосигнала (102);

10

- модуль (160) вычисления информации формант, выполненный с возможностью вычисления информации (162) формирования речевого спектра из коэффициентов (122; 322) прогнозирования;

15

- модуль (150; 350; 350'; 550) вычисления параметров усиления, выполненный с возможностью вычисления параметра (g_n ; g_c) усиления из невокализованного остаточного сигнала и информации (162) формирования спектра; и

- модуль (190; 690) формирования потоков битов, выполненный с возможностью формирования выходного сигнала (192; 692) на основе информации (142), связанной с вокализованным кадром сигнала, параметром (g_n ; g_c) усиления или параметром ($\hat{g}_c$; $\hat{g}_n$) квантованного усиления и коэффициентами (122; 322) прогнозирования.

20

2. Кодер по п. 1, причем модуль (150; 350; 350'; 550) вычисления параметров усиления содержит генератор (350a) шума, выполненный с возможностью формирования шумоподобного сигнала ($n(n)$) для кодирования; и

25

причем модуль (150; 350; 350'; 550) вычисления параметров усиления содержит процессор 350d формирования, выполненный с возможностью формирования спектра шумоподобного сигнала ($n(n)$) с использованием информации (162) формирования речевого спектра, и регулируемый усилитель (350e), выполненный с возможностью усиления спектра шумоподобного сигнала (350f) определенной формы для получения усиленного шумоподобного сигнала (350g) определенной формы, и контроллер (350k, 550n) для вычисления параметра (g_n , g_c) усиления на основании усиленного шумоподобного сигнала (350g) определенной формы.

30

3. Кодер по п. 1, дополнительно содержащий решающий модуль (130), выполненный с возможностью определения того, определен или нет остаточный сигнал из аудиокдра невокализованного сигнала.

35

4. Кодер по п. 1 или 3, в котором модуль (150; 350; 350'; 550) вычисления параметров усиления содержит:

- генератор (350a) шума, выполненный с возможностью формирования шумоподобного сигнала ($n(n)$) для кодирования;

40

- формирователь (350c), выполненный с возможностью усиления (350e) и формирования (350d) спектра шумоподобного сигнала ($n(n)$) для кодирования с использованием информации (162) формирования речевого спектра и параметра (g_n) усиления в качестве временного параметра ($g_n(\text{temp})$) усиления, чтобы получить усиленный шумоподобный сигнал (350g) для кодирования определенной формы;

45

- модуль (350h) сравнения, выполненный с возможностью сравнения невокализованного остаточного сигнала и усиленного шумоподобного сигнала (350g) для кодирования определенной формы, чтобы получить показатель для сходства между невокализованным остаточным сигналом и усиленным шумоподобным сигналом (350g)

для кодирования определенной формы; и

- контроллер (350k), выполненный с возможностью определения параметра (g_n) усиления и адаптации временного параметра ($g_n(\text{temp})$) усиления на основе результата сравнения;

5 - при этом контроллер (350k; 550n) выполнен с возможностью предоставлять параметр (g_n) усиления при кодировании в модуль формирования потоков битов, когда значение показателя для сходства выше порогового значения.

5. Кодер по п. 1 или 3, в котором модуль (150; 350; 350'; 550) вычисления параметров

10 усиления содержит:

- генератор (350a) шума, выполненный с возможностью формирования шумоподобного сигнала для кодирования;

- формирователь (350c), выполненный с возможностью усиления (350e) и формирования (350d) спектра шумоподобного сигнала ($n(n)$) для кодирования с использованием информации (162) формирования речевого спектра и параметра (g_n) усиления в качестве временного параметра ($g_n(\text{temp})$) усиления, чтобы получить усиленный шумоподобный сигнал (350g) для кодирования определенной формы;

15 - синтезатор (350m'), выполненный с возможностью синтеза синтезированного сигнала (350l') из усиленного шумоподобного сигнала (350g) для кодирования определенной формы и коэффициентов (122; 322) прогнозирования и предоставлять синтезированный сигнал (350l');
20

- модуль (350h') сравнения, выполненный с возможностью сравнения аудиосигнала (102) и синтезированного сигнала (350l'), чтобы получить показатель для сходства между аудиосигналом (102) и синтезированным сигналом (350l'); и

25 - контроллер (350k), выполненный с возможностью определения параметра (g_n) усиления и адаптации временного параметра ($g_n(\text{temp})$) усиления на основе результата сравнения;

30 - при этом контроллер (350k) выполнен с возможностью предоставлять параметр (g_n) усиления при кодировании в модуль формирования потоков битов, когда значение показателя для сходства выше порогового значения.

6. Кодер по п. 5, дополнительно содержащий запоминающее устройство (350n') усилений, выполненное с возможностью записи информации кодирования, содержащей параметр (g_n ; g_c) усиления при кодировании или информацию $\hat{g}_n$, связанную с ним, при этом контроллер (350k) выполнен с возможностью записи информации кодирования во время обработки аудиокadra и определения параметра (g_n ; g_c) усиления для последующего кадра аудиосигнала (102) на основе информации кодирования предшествующего кадра аудиосигнала (102).

40 7. Кодер по п. 4, в котором генератор (350a) шума выполнен с возможностью формирования множества случайных сигналов и комбинирования множества случайных сигналов, чтобы получить шумоподобный сигнал ($n(n)$) для кодирования.

8. Кодер по п. 1, дополнительно содержащий квантователь (170), выполненный с возможностью приема параметра (g_n ; g_c) усиления, квантования параметра (g_n ; g_c)

45 усиления, чтобы получить параметр ($\hat{g}_c$; $\hat{g}_n$) квантованного усиления.

9. Кодер по п. 4, в котором формирователь (350; 350') выполнен с возможностью комбинирования спектра шумоподобного сигнала ($n(n)$) для кодирования или спектра, извлекаемого из него, и передаточной функции ($Ffe(z)$), содержащей:

$$Ffe(z) = \frac{A(z/w1)}{A(z/w2)},$$

при этом $A(z)$ соответствует полиному фильтра кодирующего фильтра для фильтрации адаптированного шумоподобного сигнала для кодирования определенной формы, взвешенного посредством весовых коэффициентов $w1$ или $w2$, при этом $w1$ содержит положительное ненулевое скалярное значение самое большее в 1,0 и при этом $w2$ содержит положительное ненулевое скалярное значение самое большее в 1,00, при этом $w2$ превышает $w1$.

10. Кодер по п. 4, в котором формирователь (350; 350') выполнен с возможностью комбинирования спектра шумоподобного сигнала для кодирования или спектра, извлекаемого из него, с передаточной функцией ($Ft(z)$), содержащей:

$$Ft(z) = 1 - \beta z^{-1},$$

при этом z указывает представление в z -области, при этом β представляет показатель (вокализацию) для вокализации, определенной посредством связи энергии предыдущего кадра аудиосигнала и энергии текущего кадра аудиосигнала, при этом показатель β определяется в функции значения вокализации.

11. Декодер (200) для декодирования принимаемого сигнала (202), содержащего информацию, связанную с коэффициентами (122; 322) прогнозирования, причем декодер (200) содержит:

- модуль (220) вычисления информации формант, выполненный с возможностью вычисления информации (222) формирования речевого спектра из коэффициентов прогнозирования;
- генератор (240) шума, выполненный с возможностью формирования шумоподобного сигнала ($n(n)$) для декодирования;
- формирователь (250), выполненный с возможностью формирования (252) спектра шумоподобного сигнала ($n(n)$) для декодирования или его усиленного представления с использованием информации (222) формирования спектра, чтобы получать шумоподобный сигнал (258) для декодирования определенной формы; и
- синтезатор (260), выполненный с возможностью синтеза синтезированного сигнала (262) из усиленного шумоподобного сигнала (258) для кодирования определенной формы и коэффициентов (122; 322) прогнозирования.

12. Декодер по п. 11, причем формирователь (250) выполнен с возможностью формирования (252) спектра шумоподобного сигнала ($n(n)$) для декодирования с использованием информации спектрального формирования и для усиления (254) спектра шумоподобного сигнала (256) для декодирования определенной формы, чтобы получить усиленный шумоподобный сигнал (258, 1022') для декодирования определенной формы.

13. Декодер по п. 11, в котором принимаемый сигнал (202) содержит информацию, связанную с параметром (g_n ; g_c) усиления, при этом формирователь (250) содержит усилитель (254), выполненный с возможностью усиления шумоподобного сигнала ($n(n)$) для декодирования или шумоподобного сигнала (256) для декодирования определенной формы.

14. Декодер по любому из пп. 11-13, в котором принимаемый сигнал (202) дополнительно содержит вокализованную информацию (142), связанную с вокализованным кадром кодированного аудиосигнала (102), при этом декодер (200) дополнительно содержит процессор (270) вокализованных кадров, выполненный с возможностью определения вокализованного сигнала (272) на основе вокализованной информации (142), при этом декодер (200) дополнительно содержит модуль (280) комбинирования, выполненный с возможностью комбинирования синтезированного

сигнала (262) и вокализованного сигнала (272), чтобы получать кадр последовательности (282) аудиосигналов.

15. Способ (1200) для кодирования аудиосигнала (102), содержащий этапы, на которых:

- 5 - извлекают (1210) коэффициенты (122; 322) прогнозирования и остаточный сигнал из кадра (102) аудиосигнала;
- вычисляют (1220) информацию (162) формирования речевого спектра из коэффициентов (122; 322) прогнозирования;
- вычисляют (1230) параметр (g_n ; g_c) усиления из невокализованного остаточного
- 10 сигнала и информации (162) формирования спектра; и
- формируют (1240) выходной сигнал (192; 692) на основе информации (142), связанной с вокализованным кадром сигнала, параметром (g_n ; g_c) усиления или параметром ($\hat{g}_c$; $\hat{g}_n$) квантованного усиления и коэффициентами (122; 322) прогнозирования.

15 16. Способ (1300) для декодирования принимаемого аудиосигнала (202), содержащего информацию, связанную с коэффициентами прогнозирования и параметрами (g_n ; g_c) усиления, при этом способ содержит этапы, на которых:

- вычисляют (1310) информацию (222) формирования речевого спектра из коэффициентов (122; 322) прогнозирования;
- 20 - формируют (1320) шумоподобный сигнал ($p(n)$) для декодирования;
- формируют (1330) спектр шумоподобного сигнала ($p(n)$) для декодирования или его усиленное представление с использованием информации (222) формирования спектра, чтобы получать шумоподобный сигнал (258) для декодирования определенной формы;
- и
- 25 - синтезируют (1340) синтезированный сигнал (262) из усиленного шумоподобного сигнала (258) для кодирования определенной формы и коэффициентов (122; 322) прогнозирования.

17. Цифровой носитель хранения данных, имеющий сохраненную на нем компьютерную программу, которая, когда выполняется на компьютере, имеющем

30 программный код для осуществления, выполняет на компьютере способ по п. 15.

18. Цифровой носитель хранения данных, имеющий сохраненную на нем компьютерную программу, которая, когда выполняется на компьютере, имеющем программный код для осуществления, выполняет на компьютере способ по п. 16.

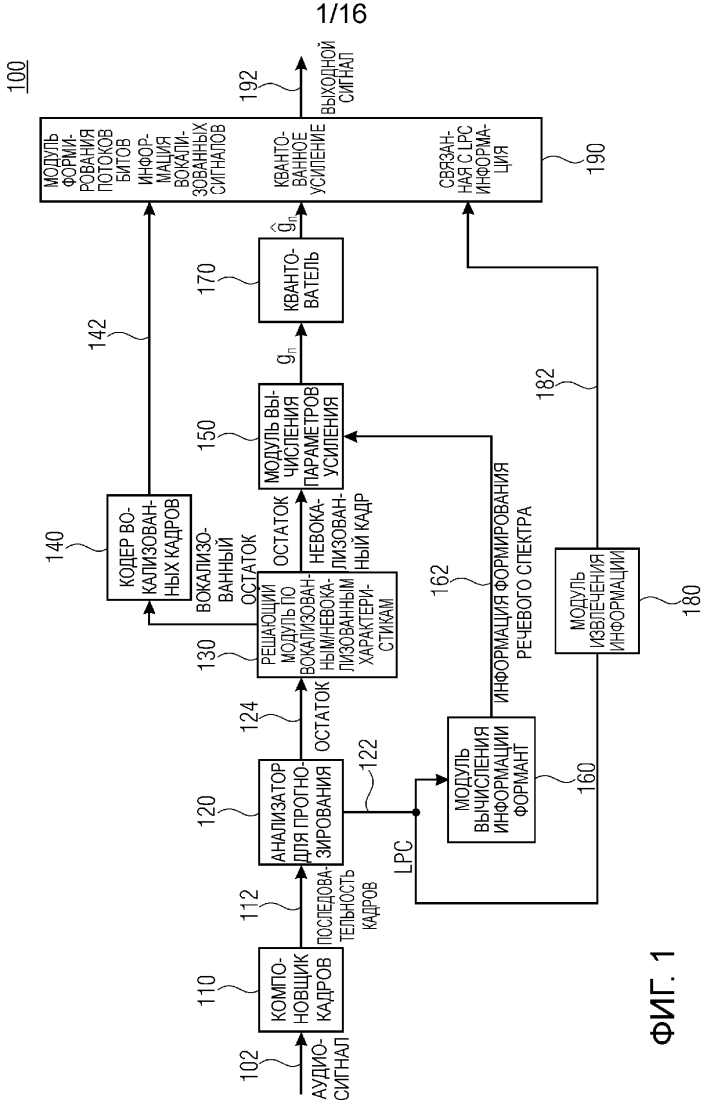
35

40

45

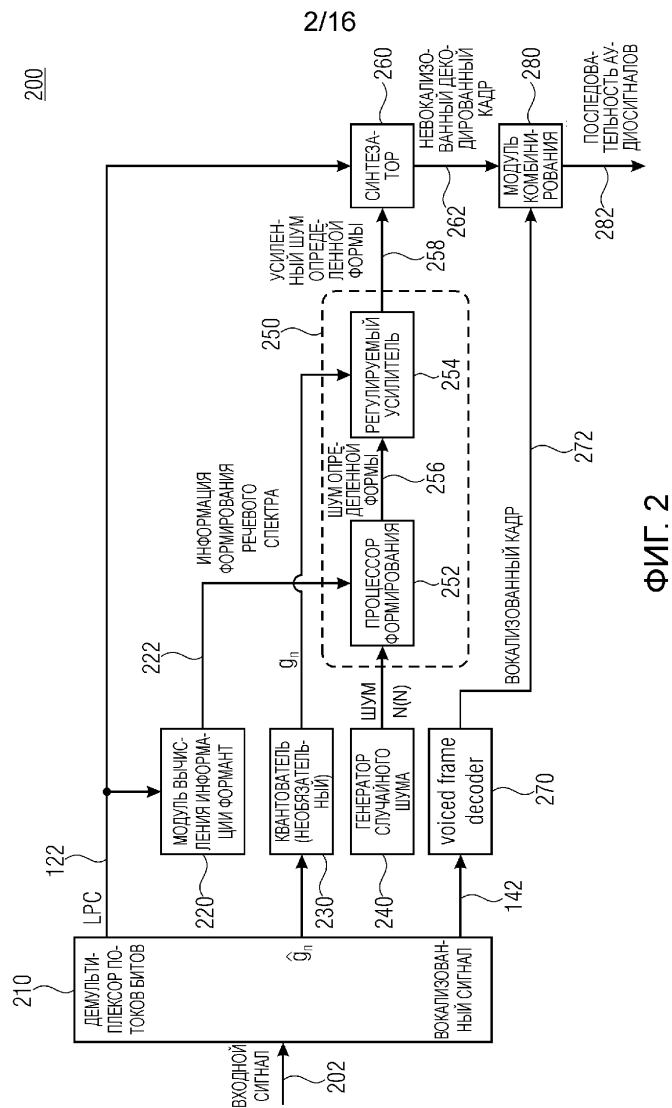
1

533451



ФИГ. 1

2

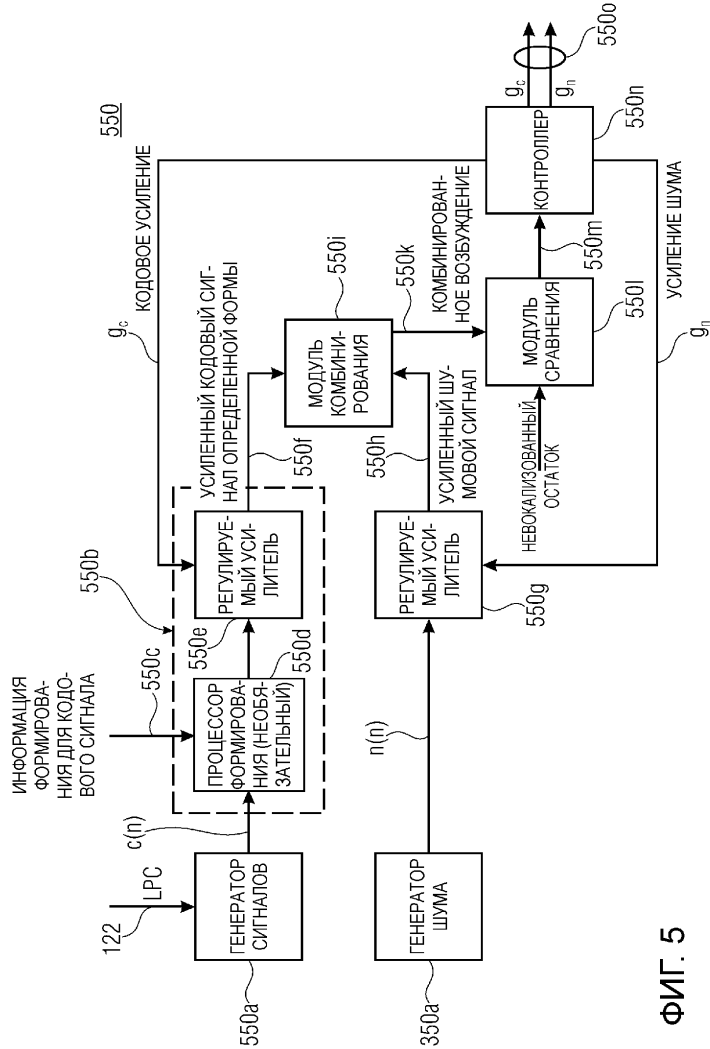


ФИГ. 2

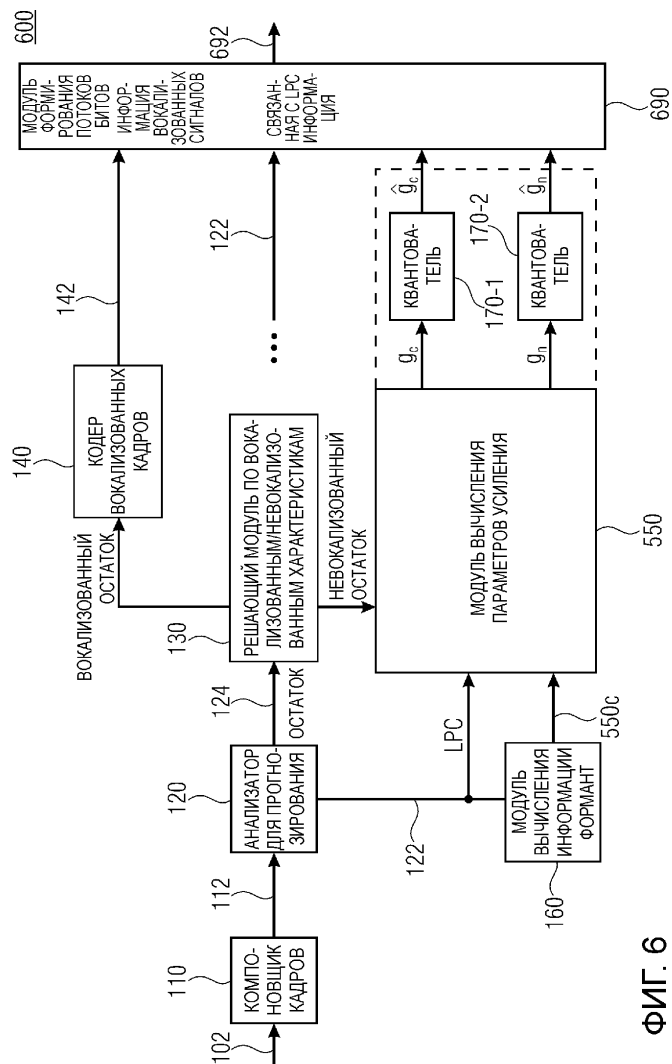
ΦΙΓ. 3

Φιν. 4

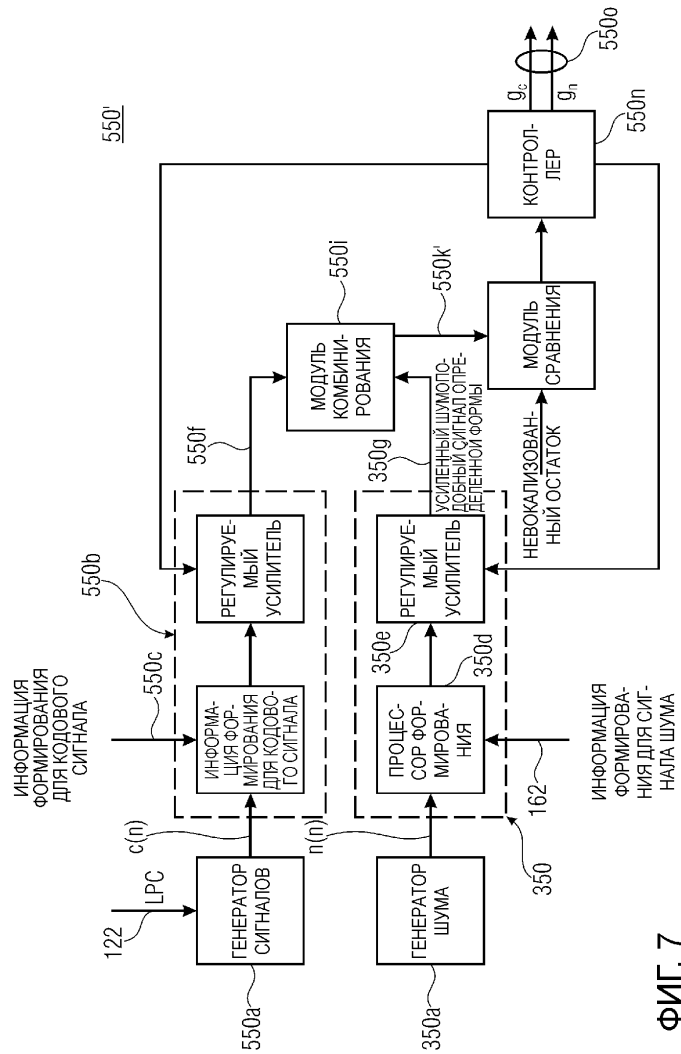




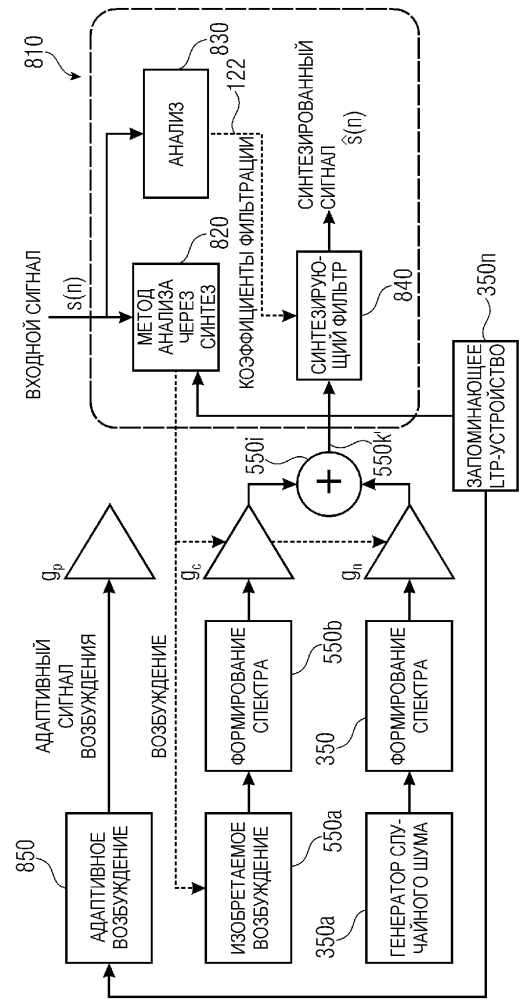
ФИГ. 5



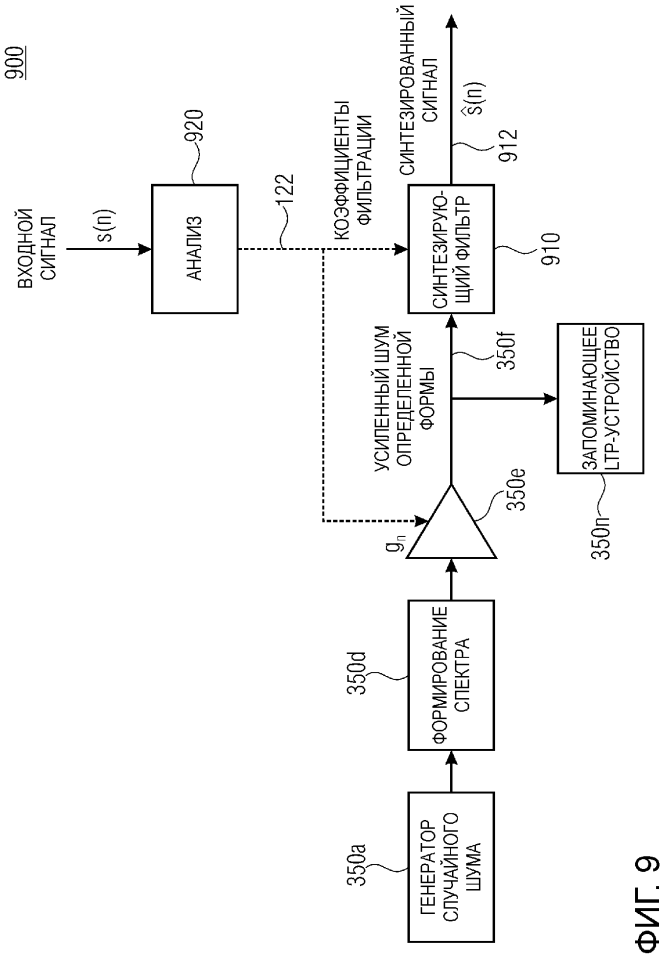
ΦΙΓ. 6



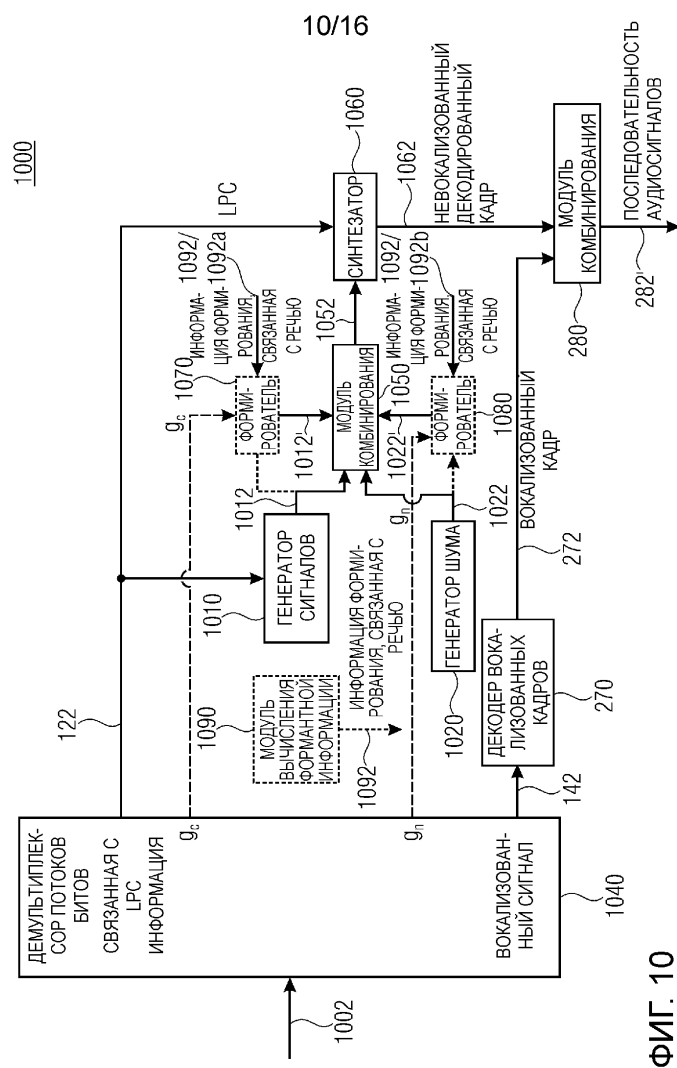
ФИГ. 7



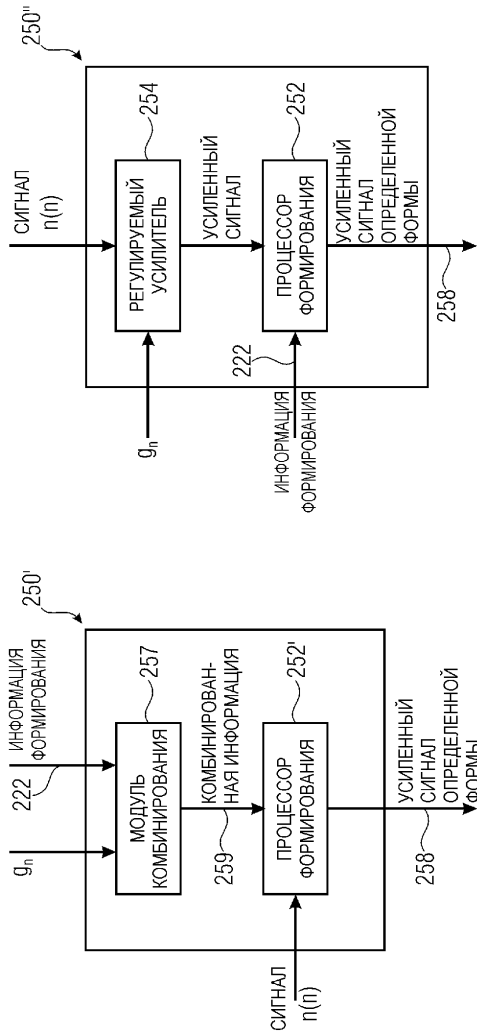
ФИГ. 8



ФИГ. 9



11/16

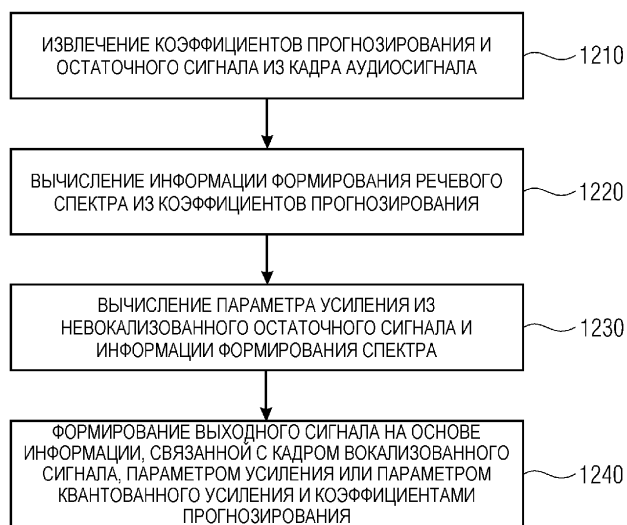


ФИГ. 11В

ФИГ. 11А

12/16

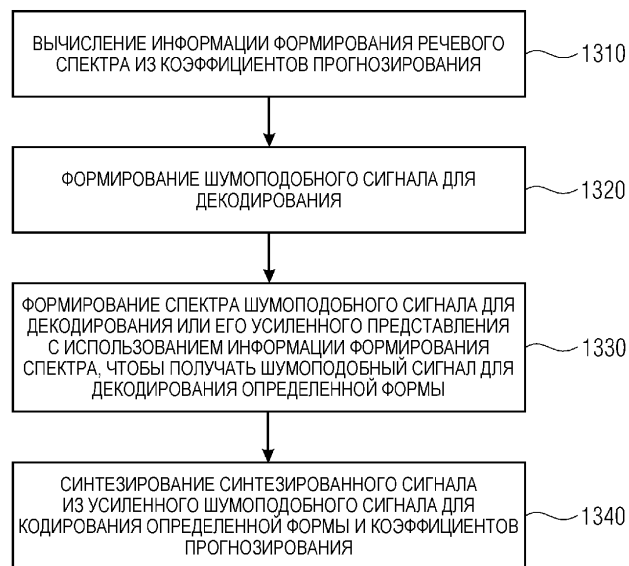
1200



ФИГ. 12

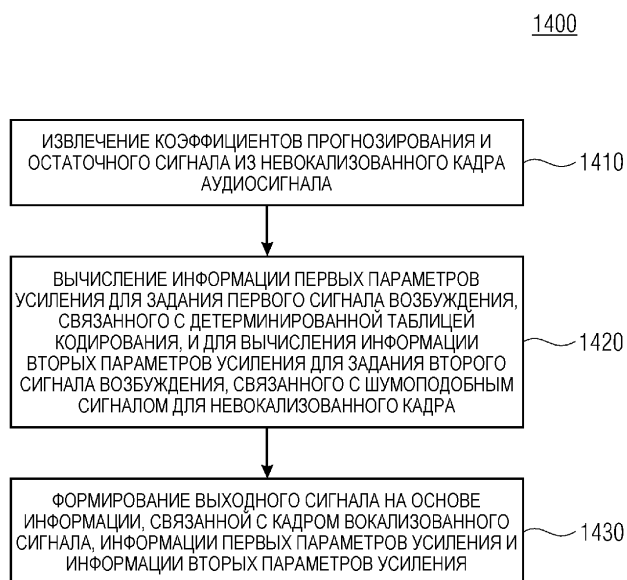
13/16

1300



ФИГ. 13

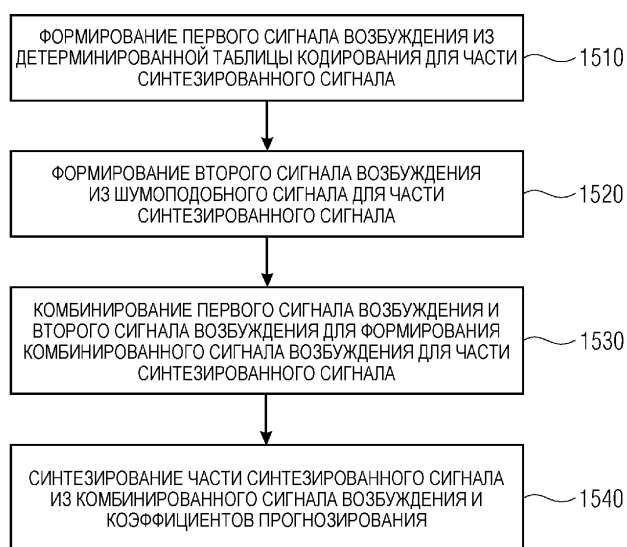
14/16



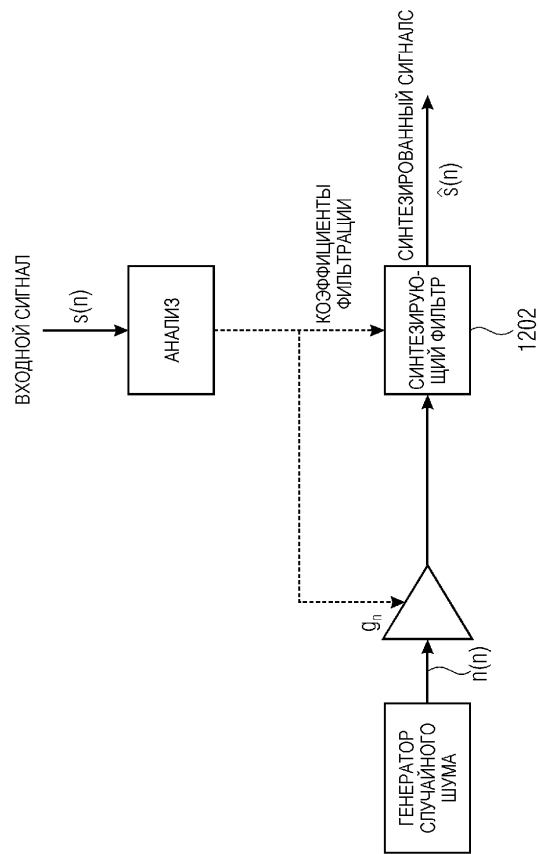
ФИГ. 14

15/16

1500



ФИГ. 15



ФИГ. 16