



ФЕДЕРАЛЬНАЯ СЛУЖБА
ПО ИНТЕЛЛЕКТУАЛЬНОЙ СОБСТВЕННОСТИ

(12) ОПИСАНИЕ ИЗОБРЕТЕНИЯ К ПАТЕНТУ

(52) СПК

G10L 19/028 (2006.01); G10L 21/038 (2006.01)

(21)(22) Заявка: 2016146738, 24.07.2015

(24) Дата начала отсчета срока действия патента:
24.07.2015

Дата регистрации:
19.09.2018

Приоритет(ы):

(30) Конвенционный приоритет:
28.07.2014 EP 14178777.0

(43) Дата публикации заявки: 28.08.2018 Бюл. № 25

(45) Опубликовано: 19.09.2018 Бюл. № 26

(85) Дата начала рассмотрения заявки РСТ на
национальной фазе: 28.02.2017

(86) Заявка РСТ:
EP 2015/067058 (24.07.2015)

(87) Публикация заявки РСТ:
WO 2016/016144 (04.02.2016)

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

(72) Автор(ы):

ДИШ Саша (DE),
ГАЙГЕР Ральф (DE),
НИДЕРМАЙЕР Андреас (DE),
НОЙЗИНГЕР Маттиас (DE),
ШМИДТ Константин (DE),
ВИЛЬДЕ Штефан (DE),
ШУБЕРТ Беньямин (DE),
НОЙКАМ Кристиан (DE)

(73) Патентообладатель(и):

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

(56) Список документов, цитированных в отчете
о поиске: US 2014/0188464 A1, 03.07.2014. US
2013/0006645 A1, 13.01.2013. US 2013/ 0346087
A1, 26.12.2013. RU 2402827 C2, 27.10.2010. RU
2381572 C2, 10.02.2010. US 8583445 B2,
12.11.2013. EP 2704142 A1, 05.03.2014. EP
2728577 A2, 07.05.2014. EP 2709106 A1,
19.03.2014. WO 2014//041020 A1, 20.03.2014.

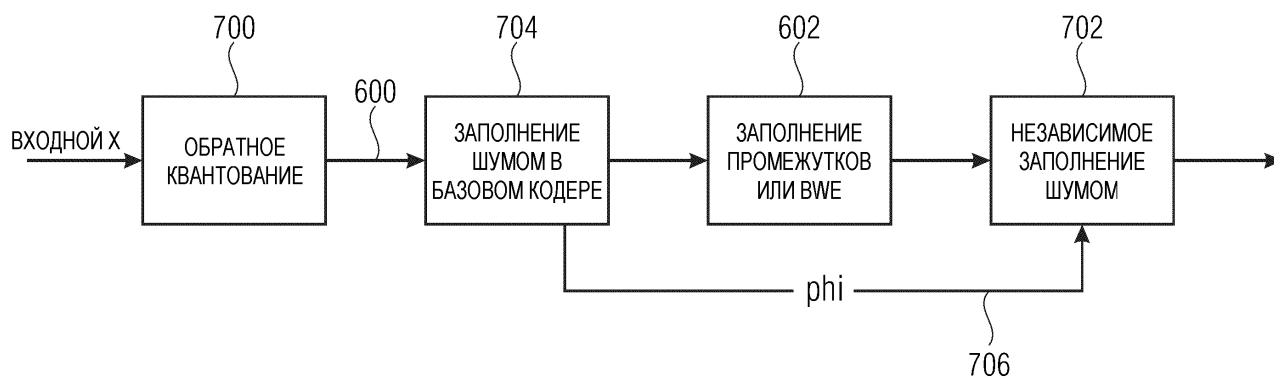
(54) УСТРОЙСТВО И СПОСОБ ФОРМИРОВАНИЯ РАСШИРЕННОГО СИГНАЛА С
ИСПОЛЬЗОВАНИЕМ ЗАПОЛНЕНИЯ НЕЗАВИСИМЫМ ШУМОМ

(57) Реферат:

Изобретение относится к обработке аудиосигнала. Технический результат – обеспечение формирования расширенного сигнала из входного сигнала. Устройство содержит: средство отображения для отображения исходного участка спектра входного сигнала на целевой участок в участке спектра расширения, исходный участок спектра содержит участок заполнения шумом; и заполнитель шумом, сконфигурированный для формирования

первых значений шума для участка заполнения шумом в исходном участке спектра входного сигнала и для формирования вторых значений шума для шумового участка в целевом участке, причем вторые значения шума являются декоррелированными от первых значений шума, или для формирования вторых значений шума для шумового участка в целевом участке, при этом вторые значения шума являются декоррелированными от первых значений шума

в исходном участке. 6 н. и 11 з.п. ф-лы, 20 ил.



ФИГ. 7

RU 2 6 6 7 3 7 6 C 2

RU 2 6 6 7 3 7 6 C 2



FEDERAL SERVICE
FOR INTELLECTUAL PROPERTY

(12) **ABSTRACT OF INVENTION**

(52) CPC

G10L 19/028 (2006.01); *G10L 21/038* (2006.01)(21)(22) Application: **2016146738, 24.07.2015**(24) Effective date for property rights:
24.07.2015Registration date:
19.09.2018

Priority:

(30) Convention priority:
28.07.2014 EP 14178777.0(43) Application published: **28.08.2018 Bull. № 25**(45) Date of publication: **19.09.2018 Bull. № 26**(85) Commencement of national phase: **28.02.2017**(86) PCT application:
EP 2015/067058 (24.07.2015)(87) PCT publication:
WO 2016/016144 (04.02.2016)Mail address:
**129090, Moskva, ul. B. Spasskaya, 25, str. 3, OOO
"Yuridicheskaya firma Gorodisskij i Partnery"**

(72) Inventor(s):

**DISCH Sascha (DE),
GEIGER Ralf (DE),
NIEDERMEIER Andreas (DE),
NEUSINGER Matthias (DE),
SCHMIDT Konstantin (DE),
WILDE Stephan (DE),
SCHUBERT Benjamin (DE),
NEUKAM Christian (DE)**

(73) Proprietor(s):

**FRAUNHOFER-GESELLSCHAFT ZUR
FOERDERUNG DER ANGEWANDTEN
FORSCHUNG E.V. (DE)**(54) **DEVICE AND METHOD OF GENERATING EXPANDED SIGNAL USING INDEPENDENT NOISE FILLING**

(57) Abstract:

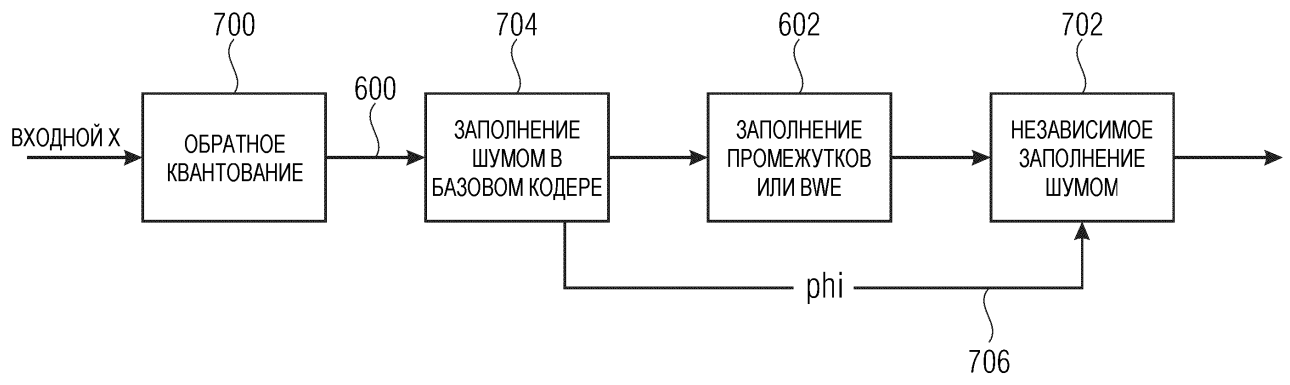
FIELD: audio equipment.

SUBSTANCE: invention relates to the processing of an audio signal. Device comprises the following: the display means for displaying the initial portion of the spectrum of the input signal to the target portion in the portion of the expansion spectrum, the initial portion of the spectrum contains the noise filling section; and the noise filler, which is configured in order to generate first noise values for the noise-filling portion in the initial portion of the spectrum of the input signal and

for generating the second noise values for the noise portion in the target portion, wherein the second noise values are decorrelated from the first noise values, or in order to generate the second noise values for the noise portion in the target portion, wherein the second noise values are decorrelated from the first noise values in the initial portion.

EFFECT: technical result is to ensure the formation of an expanded signal from the input signal.

17 cl, 20 dwg



ФИГ. 7

Изобретение относится к обработке сигналов, и конкретно - к обработке аудиосигнала.

Перцептуальное кодирование аудиосигналов с целью уменьшения объема данных для эффективного хранения или передачи этих сигналов является широко используемой практикой. В частности, когда должны обеспечиваться низкие битовые скорости передачи данных, применяемое кодирование ведет к снижению качества звучания, которое зачастую прежде всего обусловлено ограничением на стороне кодера ширины полосы аудиосигнала, подлежащего передаче. В современных кодеках существуют известные способы для восстановления сигнала на стороне декодера путем Расширения диапазона рабочих частот (BWE аудиосигнала), например, способ Копирования спектральной полосы (SBR).

В кодировании с низкой битовой скоростью часто также используется так называемое «заполнение шумом». Выступающие участки спектра, которые были квантованы в нуль в силу строгих ограничений скорости передачи данных, заполняют синтезированным шумом в декодере.

Обычно оба способа объединяют в приложениях кодирования с низкой скоростью передачи данных. Кроме того, существуют интегрированные решения, такие как Интеллектуальное заполнение промежутков (Intelligent Gap Filling, IGF), которые объединяют кодирование аудио, заполнение шумом и заполнение спектрального промежутка.

Однако все эти способы имеют общее в том, что на первом этапе соответствующий основной полосе частот (немодулированной передачи) или базовый аудиосигнал восстанавливают, используя декодирование формы сигнала и заполнение шумом, и на втором этапе выполняют обработку согласно BWE или IGF, используя полностью восстановленный сигнал. Это приводит к факту, что те же значения шума, которые были заполнены в основной полосе частот согласно заполнению шумом в ходе восстановления, используются для повторного формирования отсутствующих частей в диапазоне высоких частот (в BWE) или для заполнения оставшихся спектральных промежутков (в IGF). Использование высоко коррелированного шума для повторного создания множества участков спектра в BWE или IGF может вести к ухудшениям восприятия.

Соответствующие вопросы в уровне техники содержат:

- SBR в качестве постпроцессора для декодирования формы сигнала [1-3];
- PNS (замещение воспринимаемого шума) в AAC (усовершенствованное кодирование аудио) [4];
- заполнение шумом по стандарту USAC (объединенное кодирование речи и аудио) в MPEG-D (многоканальное кодирование) [5];
- G.719 и G.722.1C [6];
- IGF по MPEG-H 3-D [8].

Последующие публикации и заявки на патент описывают способы, которые считаются относящимися к заявке:

[1] M. Dietz, L. Liljeryd, K. Kjörling and O. Kunz, "Spectral Band Replication, a novel approach in audio coding", в 112th AES Convention, Мюнхен, Германия, 2002.

[2] S. Meltzer, R. Böhm and F. Henn, "SBR enhanced audio codecs for digital broadcasting such as "Digital Radio Mondiale" (DRM)", в 112th AES Convention, Мюнхен, Германия, 2002.

[3] T. Ziegler, A. Ehret, P. Ekstrand and M. Lutzky, "Enhancing mp3 with SBR: Features and Capabilities of the new mp3PRO Algorithm", в 112th AES Convention, Мюнхен, Германия, 2002.

[4] J. Herre, D. Schulz, Extending the MPEG-4 AAC Codec by Perceptual Noise Substitution, Audio Engineering Society 104th Convention, Предварительная публикация 4720, Амстердам, Нидерланды, 1998.

[5] Европейская заявка на патент EP2304720, заполнение шумом по стандарту USAC.

5 [6] Рекомендации G.719 и G.221C ITU-T (Международный союз электросвязи - сектор телекоммуникаций).

[7] Документ EP 2704142.

[8] Документ EP 13177350.

Аудиосигналы, обрабатываемые с помощью этих способов, страдают от артефактов, таких как грубость, модуляционные искажения и тембр, воспринимаемые как неприятные, в частности - на низкой скорости передачи данных и поэтому малой ширины полосы частот и/или появление спектральных провалов в диапазоне низких частот (LF). Причиной этого, как будет пояснено ниже, в первую очередь является факт, что восстановленные компоненты расширенного или с заполненными промежутками спектра основываются на одной или нескольких прямых копиях, содержащих шум, из основной полосы. Временные модуляции, проистекающие из упомянутой нежелательной корреляции в восстановленном шуме, являются слышимыми раздражающим образом как резкость (звука) или нежелательное искажение при восприятии. Все существующие способы, подобные mp3+SBR, AAC+SBR, USAC, G.719 и G.722.1C, а также IGF по MPEG-H 3D сначала выполняют полное базовое декодирование, включая заполнение шумом, до заполнения спектральных промежутков или диапазона высоких частот скопированными или дублированными спектральными данными из базовой полосы.

Объект настоящего изобретения состоит в обеспечении усовершенствованной идеи формирования расширенного сигнала.

25 Этот объект достигается посредством устройства для формирования расширенного сигнала по п.1, способа для формирования расширенного сигнала по п.13, системы кодирования и декодирования по п.14, способа кодирования и декодирования по п.15 или компьютерной программы по п.16.

Настоящее изобретение основано на заключении, что значительное улучшение качества звучания расширенного сигнала, сформированного путем расширения полосы частот или интеллектуального заполнения промежутков или любым другим способом формирования расширенного сигнала, имеющего спектральные значения для участка спектра расширения, не содержащиеся во входном сигнале, получают путем формирования первых значений шума для участка заполнения шумом в исходном участке спектра (источника) входного сигнала и путем формирования затем вторых независимых значений шума для шумового участка в участке назначения или целевом, то есть, в участке расширения, который теперь имеет значения шума, то есть, вторые значения шума, которые являются независимыми от первых значений шума.

40 Таким образом, устраняется проблема известного уровня техники в наличии зависимого шума в основной полосе частот и полосе частот расширения благодаря отображению спектральных значений, и устраняются связанные проблемы артефактов, таких как грубость, модуляционные искажения и тембр, воспринимаемые как неприятные в частности на низких скоростях передачи.

Другими словами, заполнение шумом вторыми значениями шума, являющимися декоррелированными от первых значений шума, то есть, значениями шума, которые, по меньшей мере, являются частично независимыми от первых значений шума, гарантирует, что артефакты более не возникают или, по меньшей мере, снижаются относительно известного уровня техники. Следовательно, соответствующая известному

уровню техники обработка спектральных значений заполнения шумом в основной полосе частот путем операции прямого расширения полосы частот или интеллектуального заполнения промежутков не декоррелирует шум от основной полосы частот, а только изменяет его уровень, например. Однако, введение декоррелированных значений шума в исходном диапазоне частот (источника) с одной стороны и в целевом диапазоне частот с другой стороны, предпочтительно выводимых на основе отдельного процесса (обработки) шума обеспечивает лучшие результаты. Однако, даже введение значений шума, не являющихся полностью декоррелированными, или не полностью независимых, но, по меньшей мере, частично декоррелированных, например, согласно значению декорреляции в 0,5 или менее, если значение декорреляции «нуль» обозначает «полностью декоррелированный», исправляет проблему полной корреляции в известном уровне техники.

Следовательно, варианты осуществления относятся к объединению декодирования формы сигнала, расширения полосы частот или заполнения промежутков и заполнения шумом в перцепционном декодере.

Дополнительные преимущества состоят в том, что в отличие от уже существующих идей, устраняются появление искажений сигнала и артефакты резкости восприятия, которые в настоящий момент являются типичными для вычисления расширений полосы частот или заполнения промежутков после декодирования формы сигнала и заполнения шумом.

Это обусловлено, в некоторых вариантах осуществления, изменением порядка упомянутых этапов обработки. Предпочтительно выполнять расширение полосы частот или заполнение промежутков непосредственно после декодирования формы сигнала, и кроме того предпочтительно вычислять заполнение шумом затем на уже восстановленном сигнале, используя некоррелированный шум.

В дополнительных вариантах осуществления декодирование формы сигнала и заполнение шумом могут выполняться в традиционном порядке и далее ниже по ходу в обработки, значения шума могут быть заменены надлежаще масштабированным некоррелированным шумом.

Следовательно, настоящее изобретение решает проблемы, которые происходят из-за операции копирования или операции дублирования (зеркального отображения) на заполненном шумом спектре, путем сдвига этапа заполнения шумом на самый конец цепочки обработки и использования некоррелированного шума для внесения вставок или заполнения промежутков.

Затем, предпочтительные варианты осуществления настоящего изобретения рассматриваются относительно сопроводительных чертежей, на которых:

Фиг.1А - иллюстрация устройства для кодирования аудиосигнала;

Фиг.1В - иллюстрация декодера для декодирования кодированного аудиосигнала, соответствующего кодеру по Фиг.1А;

Фиг.2А - иллюстрация предпочтительной реализации декодера;

Фиг.2В - иллюстрация предпочтительной реализации кодера;

Фиг.3А - иллюстрация схематичного представления спектра, как сформирован декодером спектральной области по Фиг.1В;

Фиг.3В - иллюстрация таблицы, указывающей отношение между масштабными коэффициентами для диапазонов масштабных коэффициентов и энергиями для диапазонов восстановления и информацией заполнения шумом для диапазона заполнения шумом;

Фиг.4А - иллюстрация функциональности кодера спектральной области для

применения выбора спектральных порций в первый и второй наборы спектральных порций;

Фиг.4В - иллюстрация реализации функциональности по Фиг.4А;

Фиг.5А - иллюстрация функциональности кодера с MDCT (модифицированное дискретное косинусное преобразование);

Фиг.5В - иллюстрация функциональности декодера с технологией MDCT;

Фиг.5С - иллюстрация реализации регенератора частоты;

Фиг.6 - иллюстрация блок-схемы устройства для формирования расширенного сигнала в соответствии с настоящим изобретением;

Фиг.7 - иллюстрация потока сигналов независимого заполнения шумом, управляемого информацией выбора в декодере, в соответствии с вариантом осуществления настоящего изобретения;

Фиг.8 - иллюстрация потока сигналов независимого заполнения шумом, реализованного при помощи измененного порядка следования заполнения промежутков или расширения полосы частот и заполнения шумом в декодере;

Фиг.9 - иллюстрация структурной схемы процедуры в соответствии с дополнительным вариантом осуществления настоящего изобретения;

Фиг.10 - иллюстрация структурной схемы процедуры в соответствии с дополнительным вариантом осуществления настоящего изобретения;

Фиг.11 - иллюстрация структурной схемы для пояснения масштабирования случайных значений;

Фиг.12 - иллюстрация структурной схемы, иллюстрирующей встраивание настоящего изобретения в общую процедуру расширения полосы частот или заполнения промежутков;

Фиг.13А - иллюстрация кодера с вычислением параметра расширения полосы частот; и

Фиг.13В - иллюстрация декодера с расширением полосы частот, реализованного в виде постпроцессора, а не интегрированной процедуры как на Фиг.1А или 1В.

Фиг.6 иллюстрирует устройство для формирования расширенного сигнала, такого как аудиосигнал, из входного сигнала, который также может быть аудиосигналом.

Расширенный сигнал имеет спектральные значения для участка спектра расширения, причем спектральные значения для участка спектра расширения не содержатся в первоначальном входном сигнале на входе 600 входного сигнала. Устройство содержит средство 602 отображения, чтобы отображать исходный участок спектра входного сигнала на целевой участок в участке спектра расширения, причем исходный участок спектра содержит участок заполнения шумом.

Кроме того, устройство содержит заполнитель 604 шумом, сконфигурированный для формирования первых значений шума для участка заполнения шумом в исходном участке спектра входного сигнала и для формирования вторых значений шума для шумового участка в целевом участке, причем вторые значения шума, то есть, значения шума в целевом участке, являются независимыми или некоррелированными или декоррелированными от первых значений шума в участке заполнения шумом.

Один вариант осуществления относится к ситуации, в которой заполнение шумом фактически выполняется в основной полосе частот, то есть, в которой значения шума в исходном участке были сформированы путем заполнения шумом. В дополнительной альтернативе полагают, что заполнение шумом в исходном участке не выполнялось. Однако в исходном участке имеется шумовой участок, фактически заполненный шумоподобными спектральными значениями, для примера кодированными в виде

спектральных значений исходным или базовым кодером. Отображение этого шумоподобного исходного участка на участок расширения также будет генерировать зависимый шум в исходном и целевом участках. Чтобы решить этот вопрос, заполнитель шумом только заполняет шум в целевой участок средства отображения, то есть, генерирует вторые значения шума для шумового участка в целевом участке, причем вторые значения шума являются декоррелированными от первых значений шума в исходном участке. Это замещение или заполнение шумом также может иметь место либо в буфере исходных фрагментов, либо могут иметь место непосредственно в целевом объекте. Шумовой участок может быть идентифицирован классификатором либо путем анализа исходного участка, либо путем анализа целевого участка.

С этой целью ссылка делается на Фиг.3А. Фиг.3А иллюстрирует как участок заполнения, такой как диапазон 301 масштабных коэффициентов во входном сигнале, и заполнитель шумом генерирует первые спектральные значения шума в этом диапазоне 301 заполнения шумом в операции декодирования входного сигнала.

Кроме того, этот диапазон 301 масштабных коэффициентов отображают на целевой участок, то есть, в соответствии с известным уровнем техники, сформированные значения шума отображают на целевой участок и, следовательно, целевой участок будет иметь зависимый или коррелированный с исходным участком шум.

В соответствии с настоящим изобретением, однако, заполнитель 604 шумом по Фиг.6 генерирует вторые значения шума для шумового участка в участке назначения или целевом, где вторые значения шума являются декоррелированными или некоррелированными или независимыми от первых значений шума в диапазоне 301 заполнения шумом по Фиг.3А.

В целом, заполнитель шумом и средство отображения для отображения исходного участка спектра на участок назначения могут быть включены в регенератор высокой частоты, как проиллюстрировано в контексте фигур Фиг. 1А - 5С для примера в рамках интегрированного заполнения промежутков, или могут быть реализованы в виде постпроцессора, как проиллюстрировано на Фиг.13В, и соответствующий кодер на Фиг.13А.

Обычно, входной сигнал подвергается обратному квантованию 700 или любой другой или дополнительной predetermined обработкой 700 декодера, каковое означает, что на выходе блока 700 получают входной сигнал по Фиг.6, так что вход в блок заполнения шумом базового кодера или блок 704 заполнителя шумом является входом 600 по Фиг.6. Средство отображения на Фиг.6 соответствует блоку 602 заполнения промежутков или расширения полосы частот, и блок 702 независимого заполнения шумом также включен в состав заполнителя 604 шумом по Фиг.6. Таким образом, блоки 704 и 702 оба включены в блок 604 заполнителя шумом по Фиг.6, и блок 704 генерирует так называемые первые значения шума для шумового участка в участке заполнения шумом, и блок 702 генерирует вторые значения шума для шумового участка в участке назначения или целевом, который выводят на основе участка заполнения шумом в основной полосе частот путем расширения полосы частот, выполняемого блоком 602 средства отображения или заполнения промежутков или расширения полосы частот. Кроме того, как обсуждено далее, операция независимого заполнения шумом, выполняемая блоком 702, управляется согласно вектору РНІ (параметров) управления, проиллюстрированного линией 706 управления.

1. Этап: Идентификация шума

На первом этапе идентифицируют все спектральные линии, которые представляют шум в переданном аудио кадре. Процесс идентификации может управляться согласно

уже существующим, переданным сведениям о шумовых позициях, используемых заполнением шумом [4][5], или может быть идентифицирован с помощью дополнительного классификатора. Результатом идентификации линии шума является вектор, содержащий нули и единицы, где позиция с единицей указывает спектральную

линию, которая представляет шум.

В математических терминах эта процедура может быть описана в виде:

Пусть $\hat{X} \in \mathbb{C}^N$ будет переданным и повторно квантованным спектром после заполнения шумом [4][5] кодированного с преобразованием, обрабатываемого методом окна сигнала длиной $N \in \mathbb{N}$. Пусть $m \in \mathbb{N}, 0 < m \leq N$, будет линией останова всего

процесса декодирования. Классификатор C_0 определяет спектральные линии, где используется заполнение шумом [4][5] в базовом участке:

$$C_0: \mathbb{C}^N \rightarrow \{0,1\}^m,$$

$\varphi[i] := C_0(\hat{X})[i] := \begin{cases} 1 & \text{на } \hat{X}[i] \\ 0 & \end{cases}$	использовалось заполнение шумом	, $0 \leq i < m \leq N$
	иначе	

и результат $\varphi \in \{0,1\}^m$ является вектором длины m .

Дополнительный классификатор C_1 может идентифицировать дополнительные линии в $\hat{X}$, который представляет шум. Этот классификатор может быть описан в виде:

$$C_1: \mathbb{C}^N \times \{0,1\}^m \rightarrow \{0,1\}^m,$$

$\varphi[i] := C_1(\hat{X}, \varphi)[i] := \begin{cases} 1 & \text{если} \\ 0 & \end{cases}$	$\varphi[i] = 1 \vee \hat{X}[i]$	, $0 \leq i < m \leq N$,
	классифицировано как шум	
	иначе	

После процесса идентификации шума вектор $\varphi \in \{0,1\}^m$ указателей шума задается в виде:

$\varphi[i] = \begin{cases} 1, & \text{спектральная линия} \\ 0, & \end{cases}$	$\hat{X}[i]$	, $0 \leq i < m \leq N$,
	идентифицирована как линия шума	
	спектральная линия	
	$\hat{X}[i]$	
	не идентифицирована как линия шума	

2. Этап: Независимый шум

На втором этапе конкретный участок переданного спектра выбирают и копируют в исходный фрагмент. Внутри этого исходного фрагмента идентифицированный шум заменяют случайным шумом. Энергию вставленного случайного шума настраивают на такую же энергию первоначального шума в исходном фрагменте.

В математических терминах эта процедура может быть описана в виде:

Пусть $n, n < m$, будет начальной линией для процесса копирования, описанного на этапе 3. Пусть $\hat{X}_{ST} \subset \hat{X}$ будет непрерывной частью переданного спектра $\hat{X}$, представляющей исходный фрагмент длиной $v < n$, который содержит спектральные линии $l_k, l_{k+1}, \dots, l_{k+v-1}$ в $\hat{X}$, где k - индекс первой спектральной линии в исходном

фрагменте $\hat{X}_{sT}$, так что $\hat{X}_{sT}[i] = l_{k+i}$, $0 \leq i < v$. Кроме того, пусть $\varphi' \subset \varphi$, так что $\varphi'[i] = \varphi[k+i]$, $0 \leq i < v$.

Идентифицированный шум теперь замещают сформированным случайным синтезированным шумом. Чтобы сохранить спектральную энергию на том же уровне, сначала вычисляют энергию E шума, указанного посредством φ :

$$E = \sum_{i=0}^{v-1} \varphi[i] |\hat{X}_{sT}[i]|^2.$$

Если $E = 0$ - пропустить замещение независимого шума для исходного фрагмента $\hat{X}_{sT}$, иначе заменить шум, указанный посредством φ' :

$$\hat{X}_{sT}'[i] = \begin{cases} r[i], & \varphi'[i] = 1 \\ \hat{X}_{sT}[i], & \varphi'[i] = 0 \end{cases}, \quad 0 \leq i < v,$$

где $r[i] \in \mathbb{C}$ - случайное число для всех $0 \leq i < v$.

Затем вычисляют энергию E' для вставленных случайных чисел:

$$E' = \sum_{i=0}^{v-1} \varphi[i] |\hat{X}_{sT}'[i]|^2.$$

Если $E' > 0$ - вычислить коэффициент g , иначе установить $g = 0$:

$$g = \sqrt{\frac{E}{E'}}.$$

С помощью g , повторно масштабировать замещенный шум:

$$\hat{X}_{sT}''[i] = \begin{cases} g \hat{X}_{sT}'[i], & \varphi'[i] = 1 \\ \hat{X}_{sT}'[i], & \varphi'[i] = 0 \end{cases}, \quad 0 \leq i < v.$$

После замещения шума исходный фрагмент $\hat{X}_{sT}''[i]$ содержит линии шума, которые являются независимыми от линий шума в $\hat{X}$.

3. Этап: Копирование

Исходный фрагмент $\hat{X}_{sT}''[i]$ отображают на его участок назначения в $\hat{X}$:

$$\hat{X}[c+i] = \hat{X}_{sT}''[i], \quad 0 \leq i < v, \quad c \geq n, \quad c+i < m < N,$$

или, если используется схема IGF [8]:

$$\hat{X}[c+i] = \begin{cases} \hat{X}_{sT}''[i], & \hat{X}[c+i] = 0 \\ \hat{X}[c+i], & \hat{X}[c+i] \neq 0 \end{cases}, \quad 0 \leq i < v, \quad c \geq n, \quad c+i < m < N.$$

Фиг.8 иллюстрирует вариант осуществления, в котором после какой-либо постобработки, такой как декодирование спектральной области, иллюстрируемое в блоке 112 на Фиг.1В или, в варианте осуществления постпроцессора, иллюстрируемого блоком 1326 на Фиг.13В, входной сигнал подвергают заполнению промежутков или расширению полосы частот сначала, то есть, сначала подвергают операции отображения и затем выполняют независимое заполнение шумом после этого, то есть, внутри полного спектра.

Процесс, описанный в вышеупомянутом контексте по Фиг.7, может выполняться как операция «на месте», так что промежуточный буфер $\hat{X}_{sT}''$ не является необходимым. Следовательно, порядок исполнения приспособливается (упрощается).

Исполнить первый этап как описано в контексте Фиг.7, снова набор спектральных строк $k, k+1, \dots, k+v-1$ в $\hat{X}$ является исходным участком. Выполнить:

2. Этап: Копировать

$$\hat{X}[c+i] = \hat{X}[k+i], \quad 0 \leq i < v, \quad c \geq n, \quad 0 < k+i < n, \quad c+i < m < N,$$

или, если используется схема [8] IGF:

$$\hat{X}[c+i] = \begin{cases} \hat{X}[k+i], & \hat{X}[c+i] = 0 \\ \hat{X}[c+i], & \hat{X}[c+i] \neq 0 \end{cases}$$

$$0 \leq i < v, c \geq n, \quad 0 < k+i < n, \quad c+i < m < N.$$

3. Этап: Независимое заполнение шумом

Выполнить существующее заполнение шумом вплоть до n и вычислить энергию спектральных линий шума в исходном участке $k, k+1, \dots, k+v-1$:

$$E = \sum_{i=0}^{v-1} \varphi[k+i] |\hat{X}[k+i]|^2.$$

Выполнить независимое заполнение шумом в заполнении промежутков или участка спектра BWE:

$$\hat{X}[c+i] = \begin{cases} r[i], & \varphi[k+i] = 1 \\ \hat{X}[c+i], & \varphi[k+i] = 0 \end{cases}, \quad 0 \leq i < v,$$

где $r[i]$, $0 \leq i < v$ снова является набором случайных чисел.

Вычислить энергию E' для вставленных случайных чисел:

$$E' = \sum_{i=0}^{v-1} \varphi[k+i] |\hat{X}[c+i]|^2.$$

Снова, если $E' > 0$ вычислить коэффициент g , иначе установить $g = 0$:

$$g = \sqrt{\frac{E}{E'}}.$$

С помощью g , повторно масштабировать замещенный шум:

$$\hat{X}[c+i] = \begin{cases} g\hat{X}[c+i], & \varphi[k+i] = 1 \\ \hat{X}[c+i], & \varphi[k+i] = 0 \end{cases}, \quad 0 \leq i < v.$$

Независимое заполнение шумом по изобретению может использоваться также в стерео среде пары каналов. Следовательно, кодер вычисляет надлежащее представление, L/R (правый/левый) или M/S (средний/боковой), пары каналов на каждый частотный диапазон и необязательные коэффициенты предсказания. Декодер применяет независимое заполнение шумом, как описано выше, к надлежаще выбранному представлению каналов до последующего вычисления окончательного преобразования всех частотных диапазонов в представление L/R.

Изобретение является применимым или подходящим для всех аудио приложений, в которых полная полоса частот не является доступной, или которые используют заполнение промежутков для заполнения спектральных провалов. Изобретение может найти применение в распространении или широковещании аудиоконтента такого как, например, с приложениями цифрового радио, потоковой передачи в сети Интернет и звуковой связи.

Затем, варианты осуществления настоящего изобретения обсуждаются относительно фигур Фиг. 9-12. На этапе 900 шумовые участки идентифицируют в исходном диапазоне. Эта процедура, которая была обсуждена выше по отношению к "Идентификации шума",

может основываться на дополнительной информации заполнения шумом, принятой от стороны кодера полностью, или также может быть сконфигурирована, чтобы альтернативно или дополнительно основываться на анализе сигнала относительно входного сигнала, уже сформированного, но без спектральных значений для участка спектра расширения, то есть, без спектральных значений для участка спектра этого расширения.

Затем, на этапе 902, исходный диапазон, который уже был подвергнут прямому заполнению шумом, как известно в области техники, то есть, полный исходный диапазон копируют в буфер исходных фрагментов.

Затем, на этапе 904, первые значения шума, то есть, прямые значения шума, сформированные внутри участка заполнения шумом входного сигнала, замещают в буфере исходных фрагментов случайными значениями. Затем, на этапе 906, эти случайные значения масштабируют в буфере исходных фрагментов, чтобы получить вторые значения шума для участка назначения. Затем, на этапе 908 выполняют операцию отображения, то есть, их контент буфера исходных фрагментов, доступный после этапов 904 и 906, отображают на целевой диапазон. Таким образом, посредством операции 904 замещения, и после операции 908 отображения, была получена операция независимого заполнения шумом в исходном диапазоне и в целевом диапазоне.

Фиг.10 иллюстрирует дополнительный вариант осуществления настоящего изобретения. Снова, на этапе 900, идентифицируют шум в исходном диапазоне. Однако; функциональность этого этапа 900 отличается от функциональности этапа 900 на Фиг.9, поскольку этап 900 на Фиг.9 может работать на спектре входного сигнала, который уже имеет принятые значения шума, то есть, в котором операция заполнения шумом уже была выполнена.

Однако, на Фиг.10, какая-либо операция заполнения шумом по отношению к входному сигналу не была выполнена, и входной сигнал еще не имеет каких-либо значений шума в участке заполнения шумом на входе на этапе 902. На этапе 902 исходный диапазон отображают на диапазон назначения или целевой, где значения заполнения шумом не включаются в исходный диапазон.

Таким образом, идентификация шума в исходном диапазоне на этапе 900 может быть, относительно участка заполнения шумом, выполняемая путем идентификации нулевых спектральных значений в сигнале и/или путем использования этой дополнительной информации заполнения шумом из входного сигнала, то есть, сформированной стороной кодера информации заполнения шумом. Затем, на этапе 904 считывается информация заполнения шумом и, конкретно, информация энергии, идентифицирующая энергию, подлежащая введению во входной сигнал стороны декодера.

Затем, как проиллюстрировано на этапе 1006, выполняется заполнение шумом в исходном диапазоне и, затем или одновременно, выполняется этап 1008, то есть, случайные значения вставляются в позиции в диапазоне назначения, который был идентифицирован этапом 900, по полной полосе, или которые были идентифицированы путем использования информации основной полосы частот или входного сигнала вместе с информацией отображения, то есть, какой (из множества) исходный диапазон отображен на какой (из множества) целевой диапазон.

В заключение, вставленные случайные значения масштабируют, чтобы получить вторые независимые или некоррелированные или декоррелированные значения шума.

Затем обсуждается Фиг.11 для того, чтобы проиллюстрировать дополнительную информацию о масштабировании значений заполнения шумом в участке спектра

расширения, то есть, каким образом из случайных значений получают вторые значения шума.

На этапе 1100 получают информацию энергии относительно шума в исходном диапазоне. Затем, информацию энергии определяют из случайных значений, то есть, из значений, сформированных случайным или псевдослучайным процессом, как проиллюстрировано на этапе 1102. Кроме того, этап 1104 иллюстрирует, каким образом вычислять масштабный коэффициент, то есть, путем использования информации энергии относительно шума в исходном диапазоне и путем использования информации энергии относительно случайных значений. Затем, на этапе 1106, случайные значения, то есть, на основе которых энергия была вычислена на этапе 1102, умножают на масштабный коэффициент, сформированный этапом 1104. Следовательно, иллюстрируемая на Фиг.11 процедура соответствует вычислению масштабных коэффициентов g , проиллюстрированному ранее в варианте осуществления. Однако все эти вычисления также могут выполняться в логарифмической области или в любой другой области, и этап 1106 умножения может быть заменен сложением или вычитанием в области значений по логарифмической шкале.

Дополнительная ссылка делается на Фиг.12, чтобы проиллюстрировать встраивание настоящего изобретения в рамки общей схемы интеллектуального заполнения промежутков или расширения полосы частот. На этапе 1200 информацию огибающей спектра извлекают из входного сигнала. Информация огибающей спектра может, например, формироваться выделителем 1306 параметров по Фиг.13А и может предоставляться декодером 1324 параметра по Фиг.13В. Затем, вторые значения шума и другие значения в диапазоне назначения масштабируют, используя эту информацию огибающей спектра, как проиллюстрировано на этапе 1202. Затем, любая дополнительная постобработка 1204 может выполняться для получения окончательного расширенного сигнала временной области, имеющего увеличенную полосу частот в случае расширения полосы частот или имеющего сниженное число спектральных провалов или отсутствие таковых в контексте интеллектуального заполнения промежутков.

В этом контексте в общих чертах изложено, что конкретно для варианта осуществления по Фиг.9, могут применяться несколько альтернатив. Для варианта осуществления этап 902 выполняют с полным спектром для входного сигнала или, по меньшей мере, с порцией спектра входного сигнала, которая находится выше граничной частоты заполнения шумом. Эта частота гарантирует, что ниже некоторой частоты, то есть, ниже этой частоты, какое-либо заполнение шумом не выполняют вовсе.

Затем, независимо от какой-либо специфической информации отображения исходный диапазон/целевой диапазон, полный спектр входного сигнала, то есть, весь потенциальный исходный диапазон копируют в буфер 902 исходных фрагментов и затем обрабатывают с помощью этапа 904 и 906, и этап 908 затем выбирает некоторый конкретно требуемый исходный участок из этого буфера исходных фрагментов.

В других вариантах осуществления, однако, только конкретно требуемые исходные диапазоны, которые могут быть только частями входного сигнала, копируют в одиночный буфер исходных фрагментов или в несколько отдельных буферов исходных фрагментов на основании информации исходный диапазон/целевой диапазон, включенной во входной сигнал, то есть, ассоциированной в виде дополнительной информации к этому входному сигналу аудио. В зависимости от ситуации вторая альтернатива, где только конкретно требуемые исходные диапазоны обрабатываются согласно этапам 902, 904, 906, сложность или, по меньшей мере, требования к памяти,

могут быть снижены по сравнению с ситуацией, где всегда, независимо от конкретной ситуации отображения, исходный диапазон целиком, по меньшей мере, выше граничной частоты заполнения шумом обрабатывается согласно этапам 902, 904, 906.

Затем ссылка делается на Фиг.1А - 5е для того, чтобы проиллюстрировать конкретную реализацию настоящего изобретения в регенераторе 116 частоты, который помещают до преобразователя 118 спектр-время.

Фиг.1А иллюстрирует устройство для кодирования аудиосигнала 99. Аудиосигнал 99 вводится в преобразователь 100 время-спектр для преобразования аудиосигнала, имея частоту дискретизации, в спектральное представление 101, выводимое этим преобразователем время-спектр. Спектр 101 вводится в спектральный анализатор 102 для осуществления анализа спектрального представления 101. Спектральный анализатор 101 сконфигурирован для определения первого набора первых спектральных порций 103, подлежащих кодированию с первым спектральным разрешением, и другого второго набора вторых спектральных порций 105, подлежащих кодированию со вторым спектральным разрешением. Второе спектральное разрешение меньше чем первое спектральное разрешение. Второй набор вторых спектральных порций 105 вводится в вычислитель параметра или параметрический кодер 104 для вычисления информации огибающей спектра, имеющей второе спектральное разрешение. Кроме того, кодер 106 аудио спектральной области обеспечивается для формирования первого кодированного представления 107 первого набора первых спектральных порций, имеющих первое спектральное разрешение. Кроме того, кодер вычислитель параметра/параметрический кодер 104 сконфигурирован для формирования второго кодированного представления 109 второго набора вторых спектральных порций. Первое кодированное представление 107 и второе кодированное представление 109 вводятся в мультиплексор битового потока или формирователь 108 битового потока и блок 108 в заключение выводит кодированный аудиосигнал для передачи или сохранения на запоминающем устройстве.

Обычно, первую спектральную порцию, такую как 306 по Фиг.3А, будут окружать две вторые спектральные порции, такие как 307а, 307b. Это не так в случае НЕ ААС, где частотный диапазон базового кодера является ограниченным по полосе частот.

Фиг.1В иллюстрирует декодер, соответствующий кодеру по Фиг.1А. Первое кодированное представление 107 вводится в декодер 112 аудио спектральной области, чтобы сформировать первое декодированное представление первого набора первых спектральных порций, декодированное представление имеет первое спектральное разрешение. Кроме того, второе кодированное представление 109 вводится в параметрический декодер 114, чтобы сформировать второе декодированное представление второго набора вторых спектральных порций, имеющее второе спектральное разрешение, являющееся более низким, чем первое спектральное разрешение.

Декодер дополнительно содержит регенератор 116 частоты для повторного формирования восстановленной второй спектральной порции, имеющей первое спектральное разрешение, используя первую спектральную порцию. Регенератор 116 частоты выполняет операцию заполнения фрагмента, то есть, использует фрагмент или порцию из первого набора первых спектральных порций и копирует этот первый набор первых спектральных порций в диапазон восстановления или полосу частот восстановления, имеющую вторую спектральную порцию, и обычно выполняет формирование огибающей спектра или другую операцию, как указано декодированным вторым представлением, выводимым параметрическим декодером 114, то есть, с

использованием информации второго набора вторых спектральных порций.

Декодированный первый набор первых спектральных порций и восстановленный второй набор спектральных порций, как указано на выходе регенератора 116 частоты на линии 117, вводится в преобразователь 118 спектр-время сконфигурированный для преобразования первого декодированного представления и восстановленной второй спектральной порции во временное представление 119, временное представление имеет некоторую высокую частоту дискретизации.

Фиг.2В иллюстрирует реализацию кодера по Фиг.1А. Входной аудиосигнал 99 вводится в банк 220 фильтров анализа, соответствующий преобразователю 100 время-спектр по Фиг.1А. Затем операция временного формирования шума выполняется в блоке 222 формирования временной огибающей шума (TNS). Следовательно, вход в спектральный анализатор 102 по Фиг.1А, соответствующий блоку тональной маски 226 по Фиг.2В, может быть либо полными спектральными значениями, когда операция временного формирования (огибающей) шума/временного формирования фрагмента не применяется, или могут быть значениями спектрального остатка, когда применяется операция TNS, как проиллюстрировано на Фиг.2В, блоком 222. Для двухканальных сигналов или многоканальных сигналов может дополнительно выполняться объединенное канальное кодирование 228, так что кодер 106 спектральной области по Фиг.1А может содержать блок 228 объединенного канального кодирования. Кроме того, обеспечивается энтропийный кодер 232 для выполнения сжатия данных без потерь, который является также частью кодера 106 спектральной области по Фиг.1А.

Спектральный анализатор/тональная маска (фильтр) 226 разделяет выход блока 222 TNS на базовую полосу частот и тональные компоненты, соответствующие первому набору первых спектральных порций 103, и компоненты остатка, соответствующие второму набору вторых спектральных порций 105 по Фиг.1А. Блок 224, обозначенный как кодирование выделения параметра IGF соответствует параметрическому кодеру 104 по Фиг.1А и мультиплексор битового потока 230 соответствует мультиплексору 108 битового потока по Фиг.1А.

Предпочтительно, банк 222 фильтров анализа реализован как MDCT (банк фильтров модифицированного дискретного косинусного преобразования), и MDCT используется для преобразования сигнала 99 в частотно-временную область с помощью модифицированного дискретного косинусного преобразования, действующего в качестве инструментального средства частотного анализа.

Спектральный анализатор 226 предпочтительно применяет маску тональности. Этот каскад оценки маски тональности используется, чтобы отделить тональные компоненты от шумо-подобных компонентов в сигнале. Это позволяет базовому кодеру 228 кодировать все тональные компоненты с помощью модуля психоакустики. Каскад оценки маски тональности может быть реализован многими различными способами и предпочтительно реализуется подобным по своей функциональности каскаду оценки синусоидального контура, используемому в моделировании синусоид и шума для кодирования речи/аудио [8, 9], или основанному на модели HILN кодеру аудио, описанному в [10]. Предпочтительно, используется реализация, которую легко реализовать без необходимости поддерживать траектории «смерть-гибель», но любой другой детектор тональности или шума также может использоваться.

Модуль IGF вычисляет подобие, которое существует между исходным участком и целевым участком. Целевой участок будет представлен спектром из исходного участка. Меру подобия между исходным и целевым участками создают, используя подход взаимной корреляции. Целевой участок расщепляют на $nTar$ неперекрывающихся

частотных фрагментов. Для каждого фрагмента в целевом участке создаются n_{Src} исходных фрагментов от фиксированной начальной частоты. Эти исходные фрагменты перекрываются с коэффициентом между 0 и 1, где 0 означает перекрытие 0% и 1 означает перекрытие 100%. Каждый из этих исходных фрагментов коррелирует с целевым
 5 фрагментом при различных задержках, чтобы найти исходный фрагмент, который наилучшим образом соответствует целевому фрагменту. Номер наиболее соответствующего фрагмента сохраняют в $tileNum[idx_tar]$, запаздывание, при котором он лучше всего коррелируется с целевым, сохраняют в $xcorr_lag[idx_tar][idx_src]$, и
 10 знак корреляции сохраняют в $xcorr_sign[idx_tar][idx_src]$. В случае, если корреляция является значительно отрицательной, исходный фрагмент подлежит умножению на -1 до процесса заполнения фрагмента в декодере. Модуль IGF также следит, чтобы не было наложения записи тональных компонентов в спектре, поскольку тональные компоненты сохраняют, используя маску тональности. Относящийся к полосе частот
 15 параметр энергии используется, чтобы сохранить (значение) энергии целевого участка, давая возможность восстановить спектр точно.

Этот способ имеет некоторые преимущества над классическим SBR [1] в том, что сетка гармоник многотонального сигнала сохраняется базовым кодером, тогда как только промежутки между синусоидами заполняют наиболее соответствующим
 20 "сформированным шумом" из исходного участка. Другое преимущество этой системы по сравнению с ASR (Точное спектральное замещение) [2-4] состоит в отсутствии этапа синтеза сигнала, который создает важные порции сигнала в декодере. Вместо этого эта задачу принимает на себя базовый кодер, давая возможность сохранения важных компонентов спектра. Другим преимуществом предложенной системы является
 25 непрерывная масштабируемость, которую предлагают функциональные возможности. Использование лишь $tileNum[idx_tar]$ и $xcorr_lag = 0$, для каждого фрагмента называется соответствием грубой гранулярности, и может использоваться для низких скоростей передачи, тогда как использование переменной $xcorr_lag$ для каждого
 фрагмента дает возможность лучше согласовать целевой и исходный спектры.

30 В дополнение, предлагается методика стабилизации выбора фрагмента, которая удаляет артефакты частотной области, такие как звучание с вибрацией и музыкальный шум.

В случае пар стереоканалов, применяется дополнительная обработка объединенного стерео. Это необходимо, поскольку для некоторого целевого диапазона сигнал может
 35 быть сильно коррелированным панорамируемым (с эффектом перемещения) источником звука. В случае если исходные участки, выбранные для этой конкретной области, не являются хорошо коррелированными, хотя энергии подогнаны для участков назначения, пространственное изображение может страдать из-за некоррелированных исходных участков. Кодер анализирует каждую энергетическую полосу участка назначения,
 40 обычно выполняя взаимную корреляцию спектральных значений, и если превышен некоторый порог, устанавливает флаг «объединенный» для этой энергетической полосы. В декодере энергетические полосы левого и правого канала подвергают обработке индивидуально, если этот флаг объединенного стерео не установлен. В случае если
 45 флаг объединенного стерео установлен, выполняют и энергии, и исправление в области объединенного стерео. Информацию объединенного стерео для IGF участков сигнализируют подобно информации объединенного стерео для базового кодирования, включая флаг, указывающий в случае предсказания, является ли направление предсказания от понижающего микширования к остатку, или наоборот.

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

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

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

при k , являющимся индексом частоты в области преобразования.

Другое решение состоит в том, чтобы вычислять и передавать энергии непосредственно в области объединенного стерео для диапазонов, где объединенное стерео является активным, поэтому дополнительное преобразование энергии не требуется на стороне декодера.

Исходные фрагменты всегда создаются согласно матрице Mid/Side (средний/боковой):

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

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

Регулировка энергии:

$$midTile[k] = midTile[k] \cdot midNrg[k];$$

$$sideTile[k] = sideTile[k] \cdot sideNrg[k];$$

Объединенное стерео -> преобразование LR:

Если никакой дополнительный параметр предсказания не кодируется:

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

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

Если никакой дополнительный параметр предсказания не кодируется и если сигнализируемым направлением является от «среднего» к «боковому»:

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

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

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

Если сигнализируемым направлением является от «бокового» к «среднему»:

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

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

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

Эта обработка гарантирует, что из фрагментов, используемых для повторного формирования сильно коррелированных целевых участков и панорамируемых целевых участков, результирующие левый и правый каналы все еще представляют коррелированный и панорамируемый источник звука, даже если исходные участки не являются коррелированными, сохраняя стереоизображение для таких участков.

Другими словами, в битовом потоке передаются флаги объединенного стерео, которые указывают, будет ли использоваться L/R или M/S в качестве примера для общего кодирования объединенного стерео. В декодере, во-первых, декодируют базовый сигнал, как указано флагами объединенного стерео для базовых диапазонов частот. Во-вторых, базовый сигнал сохраняют в обоих представлениях L/R и M/S. Для IGF заполнения фрагмента представление исходного фрагмента выбирают, чтобы подходило представлению целевого фрагмента, как указано информацией объединенного стерео для IGF диапазонов.

Временное формирование шума (TNS) является стандартным способом и частью усовершенствованного кодирования аудио (AAC) [11-13]. TNS можно рассматривать как расширение основной схемы перцепционного кодера введением необязательного этапа обработки между блоком фильтров и каскадом квантования. Основная задача модуля TNS состоит в том, чтобы скрыть создаваемый шум квантования во временном

участке маскирования транзиенто-подобных сигналов, и таким образом это ведет к более эффективной схеме кодирования. Во-первых, TNS вычисляет набор коэффициентов предсказания, используя "прямое предсказание" в области преобразования, например, MDCT. Эти коэффициенты затем используются для сглаживания временной огибающей сигнала. Поскольку квантование влияет на отфильтрованный спектр TNS, шум квантования тоже является временно плоским. Путем применения обратной фильтрации TNS на стороне декодера, шум квантования имеет форму в соответствии с временной огибающей фильтра TNS, и, следовательно, шум квантования маскируется транзиентом.

IGF основывается на MDCT представлении. Для эффективного кодирования должны использоваться предпочтительно длинные блоки, приблизительно в 20 мс. Если сигнал в таком длинном блоке содержит транзиенты, появляются слышимые опережающее и запаздывающее эхо в спектральных полосах IGF из-за заполнения фрагмента. Фиг.7C показывает типичный эффект опережающего эха до появления транзиента из-за IGF. Слева показана спектрограмма первоначального сигнала, и справа показана спектрограмма расширенного по полосе частот сигнала без фильтрации TNS.

Этот эффект опережающего эха снижается путем использования TNS в контексте IGF. Здесь, TNS используется в качестве инструментального средства временного формирования фрагмента, (TTS), если спектральная регенерация в декодере выполняется на остаточном сигнале TNS. Требуемые коэффициенты предсказания TTS вычисляют и применяют, используя полный спектр на стороне кодера, как обычно. На начальную и конечную частоты TNS/TTS не влияет начальная частота IGF, $f_{IGFstart}$, инструментального средства IGF. По сравнению с существующим TNS, конечная частота TTS увеличена до конечной частоты инструментального средства IGF, которая является более высокой, чем $f_{IGFstart}$. На стороне декодера коэффициенты TNS/TTS применяют на полном спектре снова, то есть, базовый спектр плюс регенерированный спектр плюс тональные компоненты из карты тональности (см. Фиг.7E). Применение TTS является необходимым для формирования временной огибающей регенерированного спектра, чтобы соответствовать огибающей исходного сигнала снова. Поэтому показанное опережающее эхо снижается. Кроме того, это все еще придает форму шуму квантования в сигнале ниже $f_{IGFstart}$, как обычно с TNS.

В действующих декодерах спектральное исправление вставкой на аудиосигнале нарушает спектральную корреляцию на границах вставки и тем самым искажает временную огибающую аудиосигнала внесением дисперсии. Следовательно, другое преимущество выполнения заполнения фрагмента IGF на остаточном сигнале состоит в том, что после применения формирующего фильтра границы фрагмента бесшовно скоррелированы, приводя к более точному временному воспроизведению сигнала.

В кодере по изобретению спектр, подвергшийся фильтрации TNS/TTS, обработке маски тональности и оценке параметра IGF, лишен какого-либо сигнала выше начальной частоты IGF, кроме тональных компонентов. Этот разреженный спектр теперь кодируется базовым кодером, использующим принципы арифметического кодирования и кодирования с предсказанием. Эти кодированные компоненты наряду с битами сигнализации образуют битовый поток аудио.

Фиг.2A иллюстрирует соответствующую реализацию декодера. Битовый поток на Фиг.2A, соответствующий кодированному аудиосигналу, вводится в демультимплексор/декодер, который будет подсоединен, относительно Фиг.1B, к блокам 112 и 114. Демультимплексор битового потока разделяет вводимый аудиосигнал на первое кодированное представление 107 по Фиг.1B и второе кодированное представление 109

по Фиг.1В. Первое кодированное представление, имеющее первый набор первых спектральных порций, вводится в блок 204 объединенного канального декодирования, соответствующий декодеру 112 участка спектра по Фиг.1b. Второе кодированное представление вводится в параметрический декодер 114, не проиллюстрированный на Фиг.2А, и затем вводится в блок 202 IGF, соответствующий регенератору 116 частоты по Фиг.1b. Первый набор первых спектральных порций, требуемых для регенерации частоты, вводится в блок 202 IGF через линию 203. Кроме того, после объединенного канального декодирования 204, специфическое базовое декодирование применяется в блоке 206 тональной маски с тем, что выход тональной маски 206 соответствует выходу декодера 112 спектральной области. Затем, выполняется объединение объединителем 208, то есть, построение кадра, где выход объединителя 208 теперь имеет спектр полного диапазона, но все еще в TNS/TTS фильтрованной области. Затем, в блоке 210, выполняют операцию обратного TNS/TTS, используя информацию фильтра TNS/TTS, предоставленную через линию 109, то есть, информация стороны TTS предпочтительно включается в первое кодированное представление, сформированное кодером 106 спектральной области, который может, например, быть непосредственным базовым кодером AAC или USAC, или также может включаться во второе кодированное представление. На выходе блока 210 обеспечивается полный спектр до максимальной частоты, которая является частотой полного диапазона, задаваемой частотой дискретизации первоначального входного сигнала. Затем, выполняется преобразование спектр/время в банке 212 фильтров синтеза, чтобы в заключение получить выходной сигнал аудио.

Фиг.3А иллюстрирует схематичное представление спектра. Спектр подразделен на диапазоны частот, SCB, масштабных коэффициентов, где имеются семь диапазонов SCB1 - SCB7 масштабных коэффициентов в иллюстрируемом примере на Фиг.3а. Диапазоны частот масштабных коэффициентов могут быть диапазонами частот масштабных коэффициентов для AAC, которые определены в стандарте AAC и имеют возрастающую ширину полосы частот к верхним частотам, как схематично проиллюстрировано на Фиг.3А. Предпочтительно выполнять интеллектуальное заполнение промежутков не с самого начала спектра, то есть, на низких частотах, а начинать операцию IGF на начальной частоте IGF, проиллюстрированной в позиции 309. Следовательно, базовый частотный диапазон простирается от самой низкой частоты до начальной частоты IGF. Выше начальной частоты IGF применяют анализ спектра, чтобы отделить спектральные компоненты 304, 305, 306, 307 с высоким разрешением (первый набор первых спектральных порций) от компонентов с низким разрешением, представленных вторым набором вторых спектральных порций. Фиг.3А иллюстрирует спектр, который является для примера входом в кодер 106 спектральной области или объединенный канальный кодер 228, то есть, базовый кодер работает в полном диапазоне, но кодирует значительное количество нулевых спектральных значений, то есть, эти нулевые спектральные значения квантованы в нуль или устанавливаются в нуль до квантования или после квантования. В любом случае, базовый кодер работает в полном диапазоне, то есть, как если спектр является таким, как проиллюстрирован, то есть, базовый декодер не обязательно должен быть осведомлен о каком-либо интеллектуальном заполнении промежутков или кодировании второго набора вторых спектральных порций с более низким спектральным разрешением.

Предпочтительно, высокое разрешение задается «по-линейным» кодированием спектральных линий, таких как линии MDCT, тогда как второе разрешение или низкое разрешение задается, например, вычислением только единственного спектрального

значения на каждый диапазон частот масштабных коэффициентов, где диапазон масштабных коэффициентов охватывает несколько частотных линий. Таким образом, второе низкое разрешение, относительно его спектрального разрешения, является

намного более низким, чем первое или высокое разрешение, определенное по-линейным кодированием, обычно применяемым базовым кодером, таким как базовый кодер AAC или USAC.

Относительно вычисления масштабного коэффициента или энергии, ситуация иллюстрируется на Фиг.3b. Вследствие того, что кодер является базовым кодером, и вследствие того, что могут иметься, но не обязательно должны иметься компоненты

первого набора спектральных порций в каждом диапазоне, базовый кодер вычисляет масштабный коэффициент для каждого диапазона частот не только в базовом диапазоне ниже начальной частоты IGF 309, но также и выше начальной частоты IGF до максимальной частоты $f_{IGFstop}$, которая меньше или равна половине частоты дискретизации, то есть, $f_s/2$. Таким образом, кодированные тональные порции 302, 304, 305, 306, 307 по Фиг.3A и, в этом варианте осуществления вместе с масштабными коэффициентами SCB1 - SCB7 соответствуют спектральным данным высокого разрешения. Спектральные данные низкого разрешения вычисляют, начиная от начальной частоты IGF, и соответствуют значениям E_1, E_2, E_3, E_4 информации энергии, которые передаются вместе с масштабными коэффициентами SF4 - SF7.

Конкретно, когда базовый кодер находится в состоянии низкой скорости передачи, дополнительная операция заполнения шумом операции в базовой полосе частот, то есть, ниже по частоте, чем начальная частота IGF, то есть, в диапазонах SCB1 - SCB3 масштабных коэффициентов вдобавок может применяться. В заполнении шумом существуют несколько смежных спектральных линий, которые были квантованы в

нуль. На стороне декодера квантованные в нуль спектральные значения повторно синтезируют, и повторно синтезированные спектральные значения корректируют по их величине, используя энергию заполнения шумом, такую как NF_2 , иллюстрируемую в позиции 308 на Фиг.3B. Энергия заполнения шумом, которая может даваться в абсолютном выражении или в относительном выражении конкретно относительно масштабного коэффициента как в USAC, соответствует энергии набора спектральных значений, квантованных в нуль. Эти спектральные линии заполнения шумом могут также рассматриваться являющимися третьим набором третьих спектральных порций, которые регенерируют синтезом непосредственного заполнения шумом без какой-либо операции IGF, основывающейся на регенерации частоты, используя частотные фрагменты из других частот, чтобы восстановить частотные фрагменты, используя спектральные значения из исходного диапазона и информацию E_1, E_2, E_3, E_4 энергии.

Предпочтительно, диапазоны, для которых вычислена информация энергии, совпадают с диапазонами масштабных коэффициентов. В других вариантах осуществления применяют группировку значений информации энергии, так что, например, для диапазонов 4 и 5 масштабных коэффициентов, передают только единственное значение информации энергии, но даже в этом варианте осуществления, границы сгруппированных диапазонов восстановлении совпадают с границами диапазонов масштабных коэффициентов. Если применяют другие разделения на диапазоны, то могут применяться некоторые повторные вычисления или вычисления синхронизации, и это может иметь смысл в зависимости от некоторой реализации.

Предпочтительно, кодер 106 спектральной области по Фиг.1A является управляемым по психоакустической модели кодером, как проиллюстрировано на Фиг.4A. Обычно,

как например, проиллюстрировано в стандарте MPEG2/4 AAC или стандарте MPEG1/2, Уровень 3, подлежащий кодированию аудиосигнал, после того, как был преобразован в спектральный диапазон (401 на Фиг.4А) пересылают на вычислитель 400 масштабного коэффициента. Вычислитель масштабного коэффициента является управляемым по

5 психоакустической модели, дополнительно принимая подлежащий квантованию аудиосигнал или принимая, как в стандарте MPEG1/2, Уровень 3 или AAC MPEG, комплексное спектральное представление аудиосигнала. Психоакустическая модель вычисляет, для каждого диапазона масштабных коэффициентов, масштабный коэффициент, представляющий психоакустический порог. Дополнительно, масштабные

10 коэффициенты затем, путем совместного действия известных внутренних и внешних итерационных циклов или путем любой другой подходящей процедуры кодирования, настраиваются так, что выполняются некоторые условия скорости передачи. Затем подлежащие квантованию спектральные значения с одной стороны и вычисленные масштабные коэффициенты с другой стороны вводятся в процессор-квантователь 404.

15 В непосредственной работе кодера аудио подлежащие квантованию спектральные значения взвешиваются согласно масштабным коэффициентам, и взвешенные спектральные значения вводятся затем в квантователь с фиксированным шагом, обычно имеющий функциональность сжатия к верхним диапазонам амплитуд. Затем, на выходе процессора-квантователя существуют индексы квантования, которые затем

20 пересылаются в энтропийный кодер, обычно имеющий конкретное и очень эффективное кодирование для набора индексов с квантованием в нуль для значений смежных частот или, как также называемых в области техники, "серии" нулевых значений.

В кодере аудио по Фиг.1А, однако, процессор-квантователь обычно принимает информацию о вторых спектральных порциях от спектрального анализатора. Таким

25 образом, процессор-квантователь 404 гарантирует, что в выходе процессора-квантователя 404 вторые спектральные порции, как идентифицировано спектральным анализатором 102, являются нулем или имеют представление, подтвержденное кодером или декодером как представление нуля, которое может быть очень эффективно закодировано, особенно когда присутствуют "серии" нулевых значений в спектре.

30 Фиг.4В иллюстрирует реализацию процессора-квантователя. Спектральные значения MDCT могут вводиться в блок 410 установки в нуль. Тогда, вторые спектральные порции уже установлены в нуль до того, как выполняется взвешивание согласно масштабным коэффициентам в блоке 412. В дополнительной реализации блок 410 не предусмотрен, но совместное действие установки в нуль выполняется в блоке 418,

35 следующем после блока 412 взвешивания. В еще дополнительной реализации операция установки в нуль может также выполняться в блоке 422 установки в нуль, следующем после квантования в блоке 420 квантователя. В этой реализации блоки 410 и 418 не будут присутствовать. Обычно, по меньшей мере, один из блоков 410, 418, 422 обеспечивается в зависимости от конкретной реализации.

40 Затем, на выходе блока 422, получают квантованный спектр, соответствующий тому, что проиллюстрировано на Фиг.3А. Этот квантованный спектр затем вводится в энтропийный кодер, такой как 232 на Фиг.2b, который может быть кодером Хаффмана или арифметическим кодером, как, например, определено в стандарте USAC.

Блоки 410, 418, 422 установки в нуль, которые обеспечиваются альтернативно друг

45 другу или параллельно, управляются спектральным анализатором 424. Спектральный анализатор предпочтительно содержит любую реализацию известного детектора тональности или содержит любой другой вид детектора, оперативно действующего для разделения спектра на компоненты, подлежащие кодированию с высоким

разрешением, и компоненты, подлежащие кодированию с низким разрешением. Другие такие алгоритмы, реализованные в спектральном анализаторе, могут быть детектором речевой активности, детектором шума, детектором речи или любым другим детектором, принимающим решение в зависимости от спектральной информации или связанных метаданных относительно требований к разрешению для различных спектральных порций.

Фиг.5А иллюстрирует предпочтительную реализацию преобразователя 100 время-спектр по Фиг.1А как, например, реализовано в AAC или USAC. Преобразователь 100 время-спектр содержит переключатель 502 временного окна, управляемый детектором 504 транзистента. Когда детектор 504 транзистента обнаруживает транзистент, то на переключатель временного окна сигнализируется переключение от длинных окон к коротким окнам. Переключатель 502 временного окна затем вычисляет, для перекрывающихся блоков, оконные кадры, где у каждого оконного кадра обычно имеются два N значений, например, 2048 значений. Затем, выполняется преобразование в преобразователе 506 блока, и этот преобразователь блока обычно дополнительно обеспечивает децимацию, так что объединенную децимацию/преобразование выполняют, чтобы получить спектральный кадр с N значениями, такими как спектральные значения MDCT. Таким образом, для операции с длинным окном, кадр на входе 506 блока содержит два N значений, например, 2048 значений, и спектральный кадр затем имеет 1024 значения. Затем однако выполняют переключение на короткие блоки, когда выполнены восемь коротких блоков, где каждый короткий блок имеет 1/8 обрабатываемых методом окна значений временной области по сравнению с длинным окном, и каждый спектральный блок имеет 1/8 спектральных значений по сравнению с длинным блоком. Таким образом, когда эту децимацию объединяют с операцией перекрытия в 50% переключателя временного окна, спектр является критически дискретизированной версией аудиосигнала 99 временной области.

Затем, ссылка делается на Фиг.5В, иллюстрирующую конкретную реализацию регенератора 116 частоты и преобразователя 118 спектр-время по Фиг.1В, или объединенную работу блоков 208, 212 по Фиг.2А. На Фиг.5В рассматривается конкретный диапазон восстановления, такой как диапазон 6 масштабного коэффициента по Фиг.3А. Первая спектральная порция в этом диапазоне частот восстановления, то есть, первая спектральная порция 306 по Фиг.3А вводится в блок 510 построителя/корректора кадров. Кроме того, восстановленная вторая спектральная порция для полосы 6 масштабного коэффициента также вводится в построитель/корректор кадров 510. Кроме того, информация энергии, такая как E_3 по Фиг.3В для полосы 6 масштабного коэффициента также вводится в блок 510. Восстановленная вторая спектральная порция в диапазоне частот восстановления уже была сформирована путем заполнения частотных фрагментов с использованием исходного диапазона, и диапазон восстановления тогда соответствует целевому диапазону. Теперь, выполняется регулировка энергии кадра, чтобы затем в заключение получить полный восстановленный кадр, имеющий N значений как, например, получено на выходе объединителя 208 по Фиг.2А. Затем, в блоке 512, выполняется обратное преобразование блока/интерполяция, чтобы получить 248 значений временной области для, например, 124 спектральных значений на входе блока 512. Затем, операция определения окон синтеза выполняется в блоке 514, который снова управляется согласно указателю длинное окно/короткое окно, переданному в виде дополнительной информации в кодированном аудиосигнале. Затем, в блоке 516, выполняется операция перекрытия/сложения с предыдущим временным кадром. Предпочтительно, MDCT применяет

перекрытие в 50% с тем, что для каждого нового временного кадра из $2N$ значений, N значений временной области в заключение выводятся. Перекрытие 50% в большой степени предпочитается вследствие того, что оно обеспечивает критическую дискретизацию и непрерывный переход от одного кадра к следующему кадру,

обусловленный операцией перекрытия/сложения в блоке 516.

Как проиллюстрировано в позиции 301 на Фиг.3А, операция заполнения шумом может дополнительно применяться не только ниже начальной частоты IGF, но также и выше начальной частоты IGF, например, для предполагаемого диапазона восстановления, совпадающей с диапазоном 6 масштабного коэффициента по Фиг.3а. Затем, спектральные значения заполнения шумом также могут вводиться в построитель/корректор 510 кадра, и корректировка спектральных значений заполнения шумом также может применяться в этом блоке, или спектральные значения заполнения шумом могут быть уже настроены с использованием энергии заполнения шумом до введения в построитель/корректор 510 кадра.

Предпочтительно, операция IGF, то есть, операция заполнения частотного фрагмента, использующая спектральные значения из других порций, может применяться в полном спектре. Таким образом, операция заполнения спектрального фрагмента может не только применяться в верхней полосе частот выше начальной частоты IGF, но также может применяться в нижней полосе частот. Кроме того, заполнение шумом без заполнения частотных фрагментов также может применяться не только ниже начальной частоты IGF, но также и выше начальной частоты IGF. Однако было установлено, что высокое качество и высокоэффективное кодирование аудио могут быть получены, когда операция заполнения шумом ограничивается частотным диапазоном ниже начальной частоты IGF, и когда операция заполнения частотных фрагментов ограничивается частотным диапазоном выше начальной частоты IGF, как проиллюстрировано на Фиг.3А.

Предпочтительно, целевые фрагменты (ТТ) (имеющие частоты более высокие, чем начальная частота IGF) привязаны к границам диапазона масштабных коэффициентов полно-скоростного кодера. Исходные фрагменты (ST), из которых берется информация, то есть, для частот более низких, чем начальная частота IGF, не привязаны к границам диапазона масштабных коэффициентов. Размер ST должен соответствовать размеру связанного ТТ. Это проиллюстрировано с использованием следующего примера. ТТ[0] имеет длину 10 элементов дискретизации MDCT. Это точно соответствует длине двух последующих SCB (например, 4+6). Затем, все возможные ST, которые подлежат корреляции с ТТ[0], также имеют длину 10 элементов дискретизации. Второй целевой фрагмент ТТ[1], являющийся смежным с ТТ[0], имеет длину 15 элементов дискретизации 1 (SCB, имеющие длину 7+8). Затем, ST для этого имеет длину 15 элементов дискретизации, а не 10 элементов дискретизации, как для ТТ[0].

Если возникает случай, что нельзя найти ТТ для ST с длиной целевого фрагмента (когда например, длина ТТ больше чем доступный исходный диапазон), то корреляцию не вычисляют, и исходный диапазон копируют несколько раз в этот ТТ (копирование выполняют одно за другим так, что частотная линия для нижней частоты второй копии следует сразу - по частоте - после частотной линии для самой высокой частоты первой копии), пока не будет полностью заполнен целевой фрагмент ТТ.

Затем делается ссылка на Фиг.5С, иллюстрирующую дополнительный предпочтительный вариант осуществления регенератора 116 частоты по Фиг.1b или блока 202 IGF по Фиг.2а. Блок 522 является генератором частотных фрагментов, принимающим не только идентификатор (ID) целевого диапазона, но и дополнительно

принимаящим ID исходного диапазона. Для примера, было определено на стороне кодера, что диапазон 3 масштабных коэффициентов по Фиг.3А очень хорошо подходит для восстановления диапазона 7 масштабных коэффициентов. Таким образом, идентификатором исходного диапазона будет 2, и идентификатором целевого диапазона будет 7. На основании этой информации, генератор 522 частотных фрагментов применяет операцию копирования или заполнения фрагмента гармониками или любую другую операцию заполнения фрагмента, чтобы сформировать необработанную вторую порцию спектральных компонентов 523. Необработанная вторая порция спектральных компонентов имеет разрешение по частоте, идентичное разрешению по частоте, включенному в первый набор первых спектральных порций.

Затем, первая спектральная порция диапазона восстановления, такая как 307 по Фиг.3А, вводится в построитель 524 кадра, и необработанная вторая порция 523 также вводится в построитель 524 кадра. Затем, восстановленный кадр корректируется корректором 526 с использованием коэффициента усиления для диапазона восстановления, вычисленного вычислителем 528 коэффициента усиления. Важно, однако, первая спектральная порция в кадре не попадает под влияние корректора 526, а только необработанная вторая часть для кадра восстановления попадает под влияние корректора 526. С этой целью, вычислитель 528 коэффициента усиления анализирует исходную полосу частот или необработанную вторую порцию 523 и дополнительно анализирует первую спектральную порцию в диапазоне частот восстановления, чтобы в заключение найти корректный коэффициент 527 усиления с тем, что энергия скорректированного кадра, выводимого корректором 526, будет энергией E_4 , когда рассматривается диапазон 7 масштабных коэффициентов.

В этом контексте очень важно оценить точность восстановления высокой частоты по настоящему изобретению по сравнению с HE-AAC. Это поясняется по отношению к диапазону 7 масштабных коэффициентов на Фиг.3А. Принимается, что кодер известного уровня техники, такой как проиллюстрирован на Фиг.13А, будет обнаруживать спектральную порцию 307, подлежащую кодированию с высоким разрешением как "отсутствующие гармоники". Затем, энергия этого спектрального компонента будет передаваться вместе с информацией спектральной огибающей для диапазона восстановления, такого как диапазон 7 масштабных коэффициентов, на декодер. Затем, декодер повторно создаст отсутствующую гармонику. Однако, спектральное значение, при котором отсутствующая гармоника 307, повторно созданная декодером известного уровня техники по Фиг.13В, будет находиться в середине диапазона 7 на частоте, обозначенной частотой 390 восстановления. Таким образом, настоящее изобретение избегает ошибки 391 по частоте, которая была бы внесена декодером известного уровня техники по Фиг.13Д.

В реализации спектральный анализатор также реализован для вычисления сходных элементов между первыми спектральными порциями и вторыми спектральными порциями и для определения на основании вычисленных сходств, для второй спектральной порции, в диапазоне восстановления первой спектральной порции, насколько возможно соответствующей второй спектральной порции. Затем, в этой реализации переменного исходного диапазона/целевого диапазона, параметрический кодер дополнительно внесет во второе кодированное представление информацию соответствия, указывающую для каждого целевого диапазона соответствующий исходный диапазон. На стороне декодера эта информация затем будет использоваться генератором 522 частотных фрагментов по Фиг.5С, иллюстрирующей формирование необработанной второй порции 523 на основании ID исходного диапазона и ID целевого

диапазона.

Кроме того, как проиллюстрировано на Фиг.3А, спектральный анализатор сконфигурирован, чтобы анализировать спектральное представление вплоть до максимальной аналитической частоты, являющейся только на небольшую величину ниже половины частоты дискретизации и предпочтительно являющейся, по меньшей мере, одной четвертой частоты дискретизации или обычно выше.

Как проиллюстрировано, кодер работает без понижающей дискретизации, и декодер работает без повышающей дискретизации. Другими словами кодер аудио спектральной области сконфигурирован для формирования спектрального представления, имеющего частоту по Найквисту, определенную частотой дискретизации первоначально введенного аудиосигнала.

Кроме того, как проиллюстрировано на Фиг.3А, спектральный анализатор сконфигурирован для анализа спектрального представления, начиная с начальной частоты заполнения промежутков, и кончая максимальной частотой, представленной максимальной частотой, включенной в спектральное представление, причем спектральная порция, простирающаяся от минимальной частоты до начальной частоты заполнения промежутков, принадлежит первому набору спектральных порций, и при этом последующие спектральные порции, такие как 304, 305, 306, 307, имеющие значения частоты выше частоты заполнения промежутков, дополнительно включаются в первый набор первых спектральных порций.

Как изложено, декодер 112 аудио спектральной области сконфигурирован так, что максимальная частота, представленная спектральным значением в первом декодированном представлении, равна максимальной частоте, включенной во временное представление, имеющее частоту дискретизации, причем спектральное значение для максимальной частоты в первом наборе первых спектральных порций является нулем или отличным от нуля. В любом случае, для этой максимальной частоты в первом наборе спектральных компонентов существует масштабный коэффициент для диапазона масштабных коэффициентов, который формируется и передается независимо от того, установлены ли в нуль или нет все спектральные значения в этом диапазоне масштабных коэффициентов, как обсуждено в контексте фигур Фиг. 3А и 3В.

Изобретение, следовательно, является полезным в том, что относительно других параметрических методик для повышения эффективности сжатия, например, замещение шума и заполнение шумом (эти способы предназначены исключительно для эффективного представления шума подобно контенту локального сигнала) изобретение позволяет точное воспроизведение частоты тональных компонентов. До настоящего времени, ни один современный способ не обращается к эффективному параметрическому представлению произвольного контента сигнала посредством заполнения спектрального промежутка, без ограничения фиксированного априорного разделения на диапазон низких частот (LF) и диапазон высоких частот (HF).

Варианты осуществления системы по изобретению улучшают современные подходы и тем самым обеспечивает высокую эффективность сжатия, отсутствие или только малое перцепционное неприятное воздействие и полную полосу частот аудиосигнала даже для низких битрейтов.

Общая система состоит из

- базового кодирования полного диапазона
- интеллектуального заполнения промежутков (заполнение фрагмента или заполнение шумом)
- разреженные тональные части в базовой полосе, выбранные тональной маской

- кодирование пары объединенного стерео для полного диапазона, включая заполнение фрагментов
- TNS на фрагменте
- спектральное «отбеливание» в диапазоне IGF

5 Первый шаг к более эффективной системе состоит в устранении необходимости для преобразования спектральных данных во вторую область преобразования, отличную от одной из базового кодера. Поскольку большинство аудиокодеков, таких как AAC, например, используют MDCT в качестве основного преобразования, является полезным выполнять BWE в области MDCT также. Вторым требованием для системы BWE будет
10 необходимость сохранять тональную сетку, посредством чего четные (даже) тональные HF компоненты сохраняются и качество кодированного аудио таким образом превосходит существующие системы. Чтобы обеспечить оба вышеупомянутых требования, была предложена система, названная Интеллектуальным заполнением промежутков (Intelligent Gap Filling, IGF). Фиг.2В показывает блок-схему предложенной
15 системы на стороне кодера, и Фиг.2А показывает систему на стороне декодера.

Затем, платформа пост-обработки описана относительно Фиг.13А и Фиг.13В, чтобы проиллюстрировать, что настоящее изобретение также может быть реализовано в восстановителе 1330 высокой частоты в этом варианте осуществления пост-обработки.

Фиг.13А иллюстрирует схематичное представление кодера аудио для технологии
20 расширения полосы частот как, например, используется в Высокоэффективном Усовершенствованном кодировании аудио (HE-AAC). Аудиосигнал в линии 1300 вводится в систему фильтров, содержащую фильтр 1302 нижних частот и фильтр 1304 высоких частот. Сигнал, выводимый фильтром 1304 высоких частот, вводится в выделитель/кодер 1306 параметров. Выделитель/кодер 1306 параметров
25 сконфигурирован для вычисления и кодирования параметров, таких как параметр спектральной огибающей, параметр добавления шума, параметр отсутствующих гармоник или параметр обратной фильтрации, например. Эти выделенные параметры вводятся в мультимплексор 1308 битового потока. Выходной сигнал пропускается через фильтр нижних частот и вводится в процессор, обычно содержащий функциональность
30 дискретизатора 1310 с пониженной частотой и базовый кодер 1312. Фильтр 1302 нижних частот ограничивает полосу частот, подлежащую кодированию, значительно меньшей полосой частот, чем имеющая место в первоначальном входном аудиосигнале на линии 1300. Это обеспечивает значительный коэффициент усиления кодирования вследствие того, что все функциональные возможности, имеющие место в базовом кодере,
35 единственно должны работать на сигнале со сниженной шириной полосой частот. Когда, например, полоса частот аудиосигнала на линии 1300 составляет 20 кГц, и когда фильтр 1302 нижних частот для примера имеет полосу частот в 4 кГц, чтобы выполнять теорему о дискретном представлении, теоретически достаточно, что сигнал, следующий после дискретизатора с пониженной частотой имеет частоту дискретизации в 8 кГц,
40 которая является значительным снижением по отношению к частоте дискретизации, требуемой для аудиосигнала 1300, которой должна быть, по меньшей мере, 40 кГц.

Фиг.13В иллюстрирует схематичное представление соответствующего декодера расширения полосы частот. Декодер содержит мультимплексор 1320 битового потока. Демультимплексор 1320 битового потока извлекает входной сигнал для базового декодера
45 1322 и входной сигнал для декодера 1324 параметра. Выходной сигнал базового декодера имеет, в вышеупомянутом примере, частоту дискретизации 8 кГц и, следовательно, полосу частот 4 кГц, тогда как для полного восстановления полосы частот выходной сигнал блока 1330 восстановления высокой частоты должен быть в 20 кГц, требующей

частоту дискретизации, по меньшей мере, 40 кГц. Чтобы сделать это возможным, требуется процессор-декодер, имеющий функциональность дискретизатора 1325 с повышением частоты, и банк 1326 фильтров. Блок 1330 восстановления высокой частоты затем принимает проанализированный по частоте низкочастотный сигнал, выводимый на банк 1326 фильтров, и восстанавливает частотный диапазон, определенный фильтром 1304 высоких частот по Фиг. 13А, используя параметрическое представление диапазона высоких частот. Блок 1330 восстановления высокой частоты имеет несколько функциональных возможностей, таких как регенерация диапазона верхних частот, использующая исходный диапазон в диапазоне низких частот, коррекция спектральной огибающей, функциональность добавления шума и функциональность для введения отсутствующих гармоник в диапазон верхних частот и, если применена и вычислена в кодере Фиг. 13А, операция обратной фильтрации, чтобы учесть факт, что более высокий частотный диапазон является обычно не столь тональным, как диапазон низких частот. В HE-AAC отсутствующие гармоники повторно синтезируют на стороне декодера и помещают точно в середине диапазона восстановления. Следовательно, все линии отсутствующих гармоник, которые были определены в некотором диапазоне частот восстановления, не были помещены в значения частоты, где они находились в первоначальном сигнале. Вместо этого эти линии отсутствующих гармоник помещены в частоты в центре некоторого диапазона. Таким образом, когда линия отсутствующей гармоник в первоначальном сигнале была помещена очень близко к границе диапазона восстановления в первоначальном сигнале, ошибка в частоте, вносимая помещением этой линии отсутствующей гармоник в восстановленный сигнал в центре диапазона, является близкой к 50% отдельного диапазона восстановления, для которого параметры были сформированы и переданы.

Кроме того, даже если типичные базовые кодеры аудио работают в спектральной области, базовый декодер все-таки формирует сигнал временной области, который затем снова преобразуют в спектральную область посредством функциональности банка 1326 фильтров. Это вносит дополнительные задержки обработки, может вносить артефакты из-за тандемной обработки сначала преобразования из спектральной области в частотную область и снова преобразования обычно в другую частотную область и, конечно, это также требует значительного объема вычислительной сложности и тем самым - электрической мощности, каковое является определенно проблемой, когда технология расширения полосы частот применяется в мобильных устройствах, таких как мобильные телефоны, планшетные компьютеры или ноутбуки, и т.д.

Хотя некоторые аспекты были описаны в контексте устройства для кодирования или декодирования, понятно, эти аспекты также представляют описание соответствующего способа, где блок или устройство соответствуют этапу способа или функции этапа способа. Аналогично, аспекты, описанные в контексте этапа способа, также представляют описание соответствующего блока или фрагмента или признака соответствующего устройства. Некоторые или все из этапов способа могут исполняться посредством (или с использованием) аппаратно реализованного устройства, подобного, например, микропроцессору, программируемому компьютеру или электронной схеме. В некоторых вариантах осуществления некоторый один или несколько из наиболее важных этапов способа могут исполняться таким устройством.

В зависимости от некоторых требований к реализации варианты осуществления изобретения могут быть реализованы в виде аппаратных средств или программного обеспечения. Реализация может выполняться с использованием не-временного носителя данных, такого как цифровой носитель данных, например, гибкий диск, накопитель на

жестком диске (HDD), цифровой многофункциональный диск (DVD), диск по технологии Blu-Ray, компакт-диск (CD), постоянное запоминающее устройство (ROM), программируемое постоянное запоминающее устройство (PROM) и стираемое программируемое постоянное запоминающее устройство (EPROM), электрически-стираемое программируемое ПЗУ (EEPROM) или флэш-память с наличием электронно-читаемых управляющих сигналов, сохраненных на них, которые взаимодействуют (или способны к взаимодействию) с программируемой компьютерной системой с тем, что выполняется соответствующий способ. Следовательно, цифровой носитель данных может быть читаемым компьютером.

Некоторые варианты осуществления согласно изобретению содержат носитель информации, имеющий электронно-читаемые управляющие сигналы, которые способны к взаимодействию с программируемой компьютерной системой с тем, что выполняется один из способов, описанных здесь.

Обычно, варианты осуществления настоящего изобретения могут быть реализованы в виде компьютерного программного продукта с кодом программы, код программы является используемым для выполнения одного из способов, когда компьютерный программный продукт работает на компьютере. Код программы может, например, сохраняться на машиночитаемом носителе.

Другие варианты осуществления содержат сохраненную на машиночитаемом носителе компьютерную программу для выполнения одного из описанных здесь способов.

Другими словами, вариантом осуществления способа по изобретению является, следовательно, компьютерная программа с наличием кода программы для выполнения одного из способов, описанных здесь, когда компьютерная программа работает на компьютере.

Дополнительным вариантом осуществления способа по изобретению является, следовательно, носитель информации (или цифровой носитель данных, или читаемый компьютером носитель данных), содержащий, записанную на нем компьютерную программу для выполнения одного из способов, описанных здесь. Носитель информации, цифровой носитель данных или носитель с записанными данными являются обычно материальными и/или не-временными.

Дополнительным вариантом осуществления способа изобретения является, следовательно, поток данных или последовательность сигналов, представляющих компьютерную программу для выполнения одного из способов, описанных здесь.

Поток данных или последовательность сигналов, например, могут быть сконфигурированы, чтобы передаваться через соединение передачи данных, например, через сеть Интернет.

Дополнительный вариант осуществления содержит средство обработки, например, компьютер или программируемое логическое устройство, сконфигурированное или приспособленное для выполнения одного из способов, описанных здесь.

Дополнительный вариант осуществления содержит компьютер с наличием установленной на нем компьютерной программы для выполнения одного из способов, описанных здесь.

Дополнительный вариант осуществления согласно изобретению содержит устройство или систему, сконфигурированную, чтобы передавать (например, электронно или оптически) на приемник компьютерную программу для выполнения одного из способов, описанных здесь. Приемник может, например, быть компьютером, мобильным устройством, запоминающим устройством и т.п. Устройство или система могут, например, содержать файловый сервер для осуществления передачи компьютерной

программы на приемник.

В некоторых вариантах осуществления программируемое логическое устройство (например, программируемая вентильная матрица) может использоваться, чтобы выполнять некоторые или все из функциональных возможностей способов, описанных здесь. В некоторых вариантах осуществления программируемая вентильная матрица может совместно действовать с микропроцессором для того, чтобы выполнять один из способов, описанных здесь. Обычно, способы предпочтительно выполняются каким-либо аппаратно реализованным устройством.

Вышеописанные варианты осуществления являются просто иллюстративными относительно принципов настоящего изобретения. Понятно, что модификации и изменения компоновок и деталей, описанных здесь, будут очевидны специалистам в данной области техники. Намерение, следовательно, состоит в том, чтобы ограничиваться только объемом последующих пунктов формулы изобретения, а не конкретными деталями, представленными с помощью описания и пояснения вариантов осуществления в настоящем документе.

(57) Формула изобретения

1. Устройство для формирования расширенного сигнала из входного сигнала (600), причем расширенный сигнал имеет спектральные значения для участка спектра расширения, при этом спектральные значения для участков спектра расширения не содержатся во входном сигнале (600), содержащее:

средство (602) отображения для отображения исходного участка спектра входного сигнала на целевой участок в участке спектра расширения, причем исходный участок спектра содержит участок (302) заполнения шумом; и

заполнитель (604) шумом, выполненный с возможностью формировать первые значения шума для участка (302) заполнения шумом в исходном участке спектра входного сигнала и формировать вторые значения шума для шумового участка в целевом участке, причем вторые значения шума являются декоррелированными от первых значений шума, или формировать вторые значения шума для шумового участка в целевом участке, причем вторые значения шума являются декоррелированными от первых значений шума в исходном участке,

при этом заполнитель (604) шумом выполнен с возможностью:

идентифицировать участок заполнения шумом, имеющий первые значения шума, во входном сигнале;

копировать (902) по меньшей мере участок входного сигнала в буфер исходных фрагментов, каковой участок содержит исходный участок спектра;

замещать (904) идентифицированные первые значения шума независимыми значениями шума; и

при этом средство (602) отображения выполнено с возможностью отображать буфер исходных фрагментов, имеющий декоррелированные значения шума, на целевой участок.

2. Устройство по п.1, при этом входной сигнал является кодированным сигналом, содержащим параметры заполнения шумом для исходного участка спектра входного сигнала, причем заполнитель шумом выполнен с возможностью формировать первые значения шума, используя параметры заполнения шумом, и формировать вторые значения шума, используя информацию энергии в отношении первых значений шума.

3. Устройство по п.1, в котором заполнитель (604) шумом выполнен с возможностью определять с помощью измерений информацию энергии в отношении декоррелированных значений (1102) шума и информацию энергии в отношении первых

значений (1100) шума и масштабировать (906) декоррелированные значения шума с использованием значения масштабирования, полученного из информации энергии в отношении декоррелированных значений шума и информации энергии в отношении первых значений шума.

5 4. Устройство по п.1, в котором заполнитель (604) шумом выполнен с возможностью формировать вторые значения шума после операции средства (602) отображения или формировать (604) первые и вторые значения шума после операции средства (602) отображения.

10 5. Устройство по п.1, в котором средство (602) отображения выполнено с возможностью отображать исходный участок спектра на целевой участок, при этом заполнитель (604) шумом выполнен с возможностью осуществлять заполнение шумом в участках спектра путем формирования первых значений шума, используя заполнение шумом и параметры заполнения шумом, переданные во входном сигнале в виде дополнительной информации, и осуществлять заполнение шумом в целевом участке 15 для формирования вторых спектральных значений с использованием информации энергии в отношении первых значений шума.

6. Устройство по п.1, дополнительно содержащее средство коррекции огибающей для коррекции (1202) вторых значений шума в участке спектра расширения с использованием информации огибающей спектра, включенной во входной сигнал в 20 виде дополнительной информации.

7. Устройство по п.1, в котором заполнитель (604) шумом выполнен с возможностью использовать только дополнительную информацию входного сигнала для идентификации спектральных позиций для заполнения шумом, или 25 заполнитель (604) шумом выполнен с возможностью анализировать временную или спектральную характеристику входного сигнала со спектральными значениями или без таковых в участке заполнения шумом, чтобы идентифицировать спектральные позиции для заполнения шумом.

8. Устройство по п.1, в котором заполнитель (604) шумом выполнен с возможностью 30 идентифицировать позиции шума с использованием вектора (706) идентификации, в котором имеются записи о спектральных позициях только в исходном спектральном участке или имеются записи о спектральных позициях в исходном участке спектра и в целевом участке.

9. Устройство по п.1, в котором 35 заполнитель (604) шумом выполнен с возможностью копировать, в операции (902) копирования, полную спектральную порцию входного сигнала или полную спектральную порцию входного сигнала выше граничной частоты заполнения шумом, обычно используемой средством отображения, в буфер исходных фрагментов и выполнять операцию (904) замещения в отношении полного буфера исходных 40 фрагментов, или

заполнитель (604) шумом выполнен с возможностью копировать, в операции (902) копирования, только участок спектра входного сигнала, идентифицированный одним или более конкретными исходными идентификаторами для исходного участка, который должен использоваться средством (602) отображения для целевого участка, причем 45 отдельный буфер исходных фрагментов используется для каждой отличной от других отдельной операции отображения.

10. Устройство по п.1, в котором средство (602) отображения выполнено с возможностью осуществлять операцию заполнения промежутков для формирования

целевого участка, при этом устройство дополнительно содержит:

аудиодекодер (112) спектральной области для формирования первого декодированного представления первого набора первых спектральных порций, причем декодированное представление имеет первое спектральное разрешение;

5 параметрический декодер (114) для формирования второго декодированного представления второго набора вторых спектральных порций, имеющих второе спектральное разрешение, являющееся более низким, чем первое спектральное разрешение;

10 регенератор (116) частоты для повторного формирования восстановленной второй спектральной порции, имеющей первое спектральное разрешение, с использованием первой спектральной порции и информации огибающей спектра для восстановленной второй спектральной порции; и

преобразователь (118) спектр-время для преобразования первого декодированного представления и восстановленной второй спектральной порции во временное
15 представление,

причем средство (602) отображения и заполнитель (604) шумом по меньшей мере частично включены в состав регенератора (116) частоты.

11. Устройство по п.10, в котором аудиодекодер спектральной области выполнен с возможностью вывода последовательности декодированных кадров спектральных
20 значений, при этом декодированный кадр является первым декодированным представлением, причем декодированный кадр содержит спектральные значения для первого набора спектральных порций и указатели нулей для второго набора вторых спектральных порций,

при этом устройство дополнительно содержит объединитель (208) для объединения
25 спектральных значений, сформированных регенератором частоты для второго набора вторых спектральных порций, и спектральных значений первого набора первых спектральных порций в диапазоне частот восстановления, чтобы получить восстановленный спектральный кадр, содержащий спектральные значения для первого набора первых спектральных порций и второго набора вторых спектральных порций;
30 и

при этом преобразователь (118) спектр-время выполнен с возможностью преобразования восстановленного спектрального кадра во временное представление.

12. Устройство по п.1, дополнительно содержащее: для каждого целевого участка, идентификацию исходного участка, при этом средство (602) отображения выполнено
35 с возможностью выбирать исходный участок, используя идентификацию исходного участка, и отображать выбранный исходный участок на целевой участок.

13. Способ формирования расширенного сигнала из входного сигнала (600), при этом расширенный сигнал имеет спектральные значения для участка спектра расширения, причем спектральные значения для участков спектра расширения не
40 содержатся во входном сигнале (600), содержащий этапы, на которых:

отображают (602) исходный участок спектра входного сигнала на целевой участок в участке спектра расширения, причем исходный участок спектра содержит участок (302) заполнения шумом; и

формируют (604) первые значения шума для участка (302) заполнения шумом в
45 исходном участке спектра входного сигнала и формируют вторые значения шума для шумового участка в целевом участке, причем вторые значения шума являются декоррелированными от первых значений шума, или формируют вторые значения шума для шумового участка в целевом участке, при этом вторые значения шума являются

декоррелированными от первых значений шума в исходном участке,
причем формирование (604) содержит этапы, на которых:
идентифицируют участок заполнения шумом, имеющий первые значения шума, во
входном сигнале;

- 5 копируют (902) по меньшей мере участок входного сигнала в буфер исходных
фрагментов, каковой участок содержит исходный участок спектра; и
замещают (904) идентифицированные первые значения шума независимыми
значениями шума; и
при этом отображение (602) содержит отображение буфера исходных фрагментов,
10 имеющего декоррелированные значения шума, на целевой участок.

14. Система для обработки аудиосигнала, содержащая:
кодер для формирования кодированного сигнала; и
устройство для формирования расширенного сигнала по п.1,
при этом кодированный сигнал подвергается обработке (700), чтобы сформировать
15 входной сигнал (600) в устройство для формирования расширенного сигнала.

15. Способ обработки аудиосигнала, содержащий:
формирование кодированного сигнала из входного сигнала; и
способ формирования расширенного сигнала по п.13,
при этом кодированный сигнал подвергают заранее определенной обработке (700)
20 для того, чтобы сформировать входной сигнал (600) в способ формирования
расширенного сигнала.

16. Машиночитаемый носитель информации, на котором сохранен код компьютерной
программы, который при его исполнении на компьютере предписывает компьютеру
выполнять способ по п.13.

- 25 17. Машиночитаемый носитель информации, на котором сохранен код компьютерной
программы, который при его исполнении на компьютере предписывает компьютеру
выполнять способ по п.15.

30

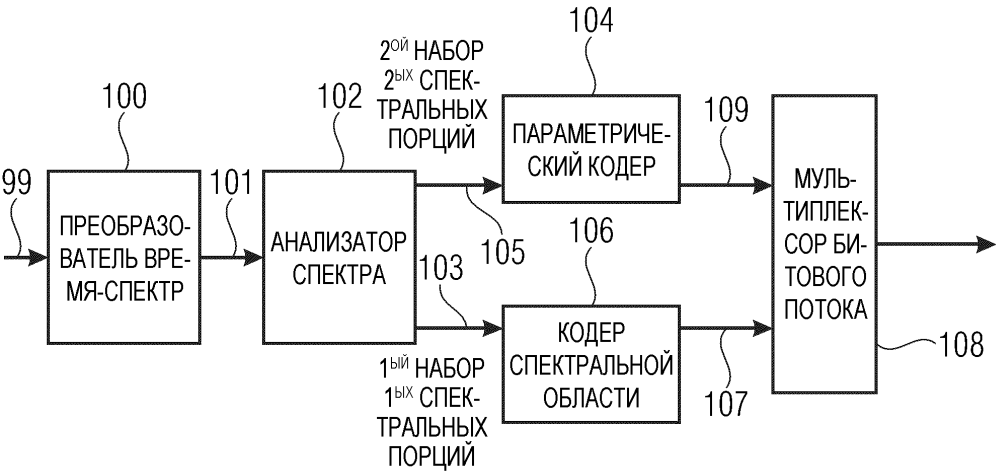
35

40

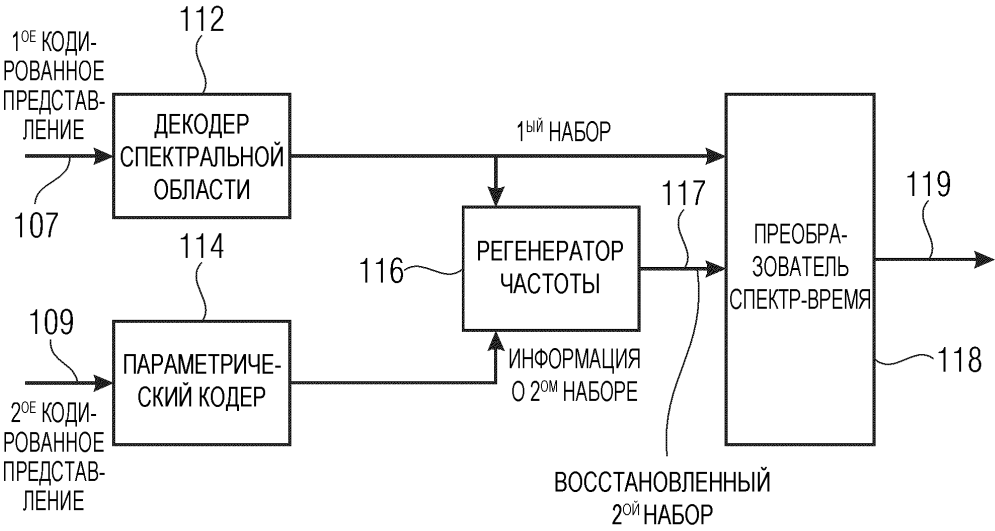
45

538259

1/17

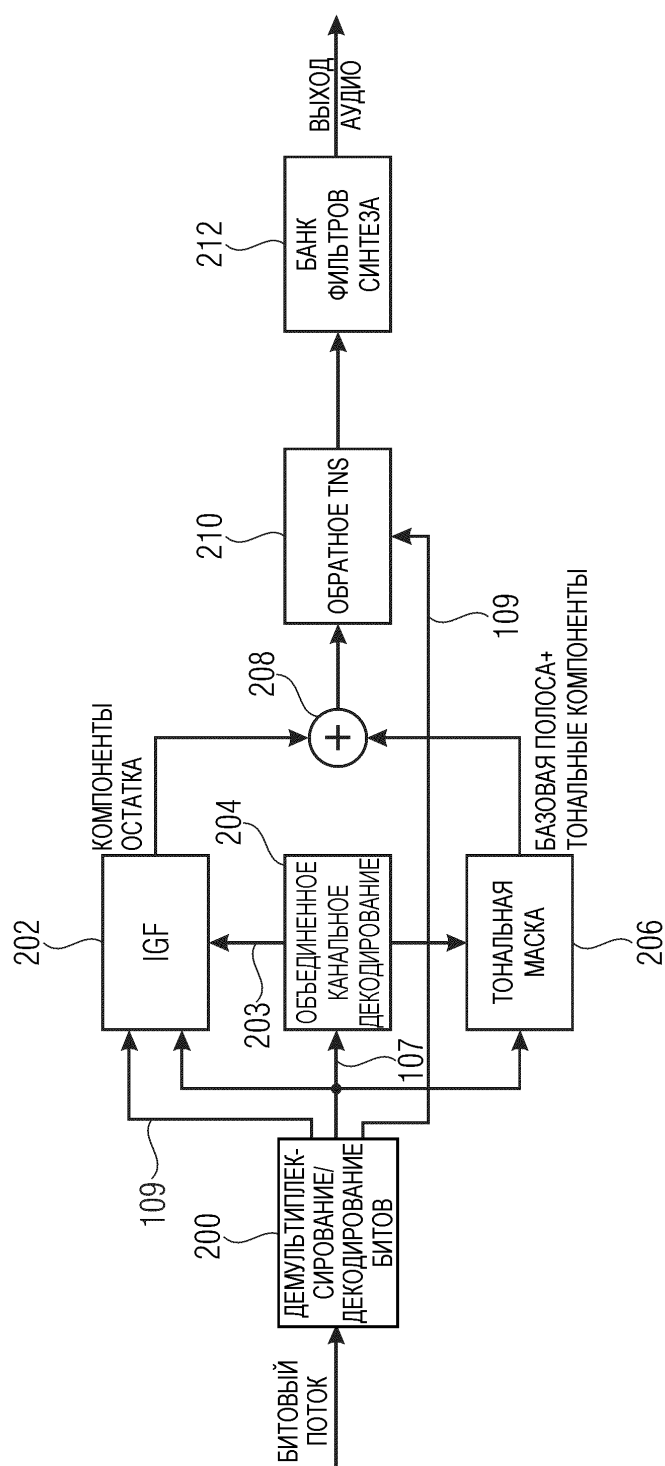


ФИГ. 1А

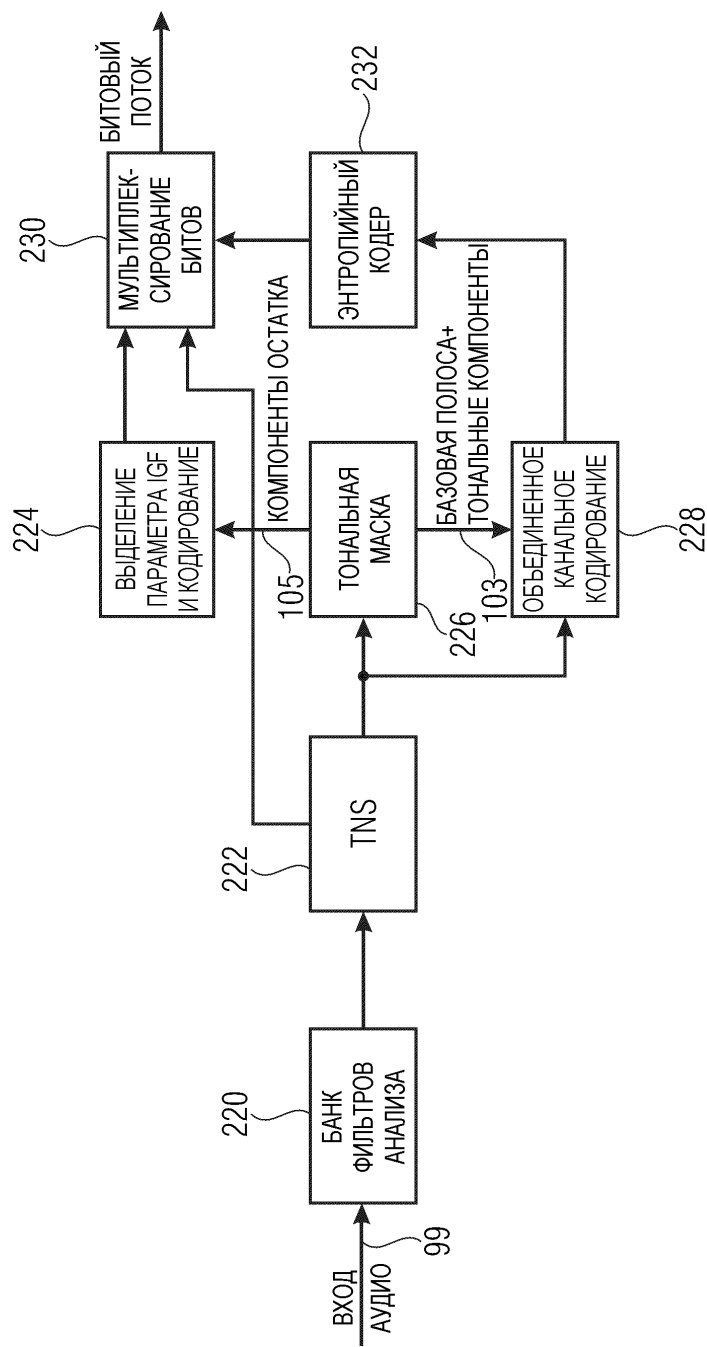


ФИГ. 1В

2/17



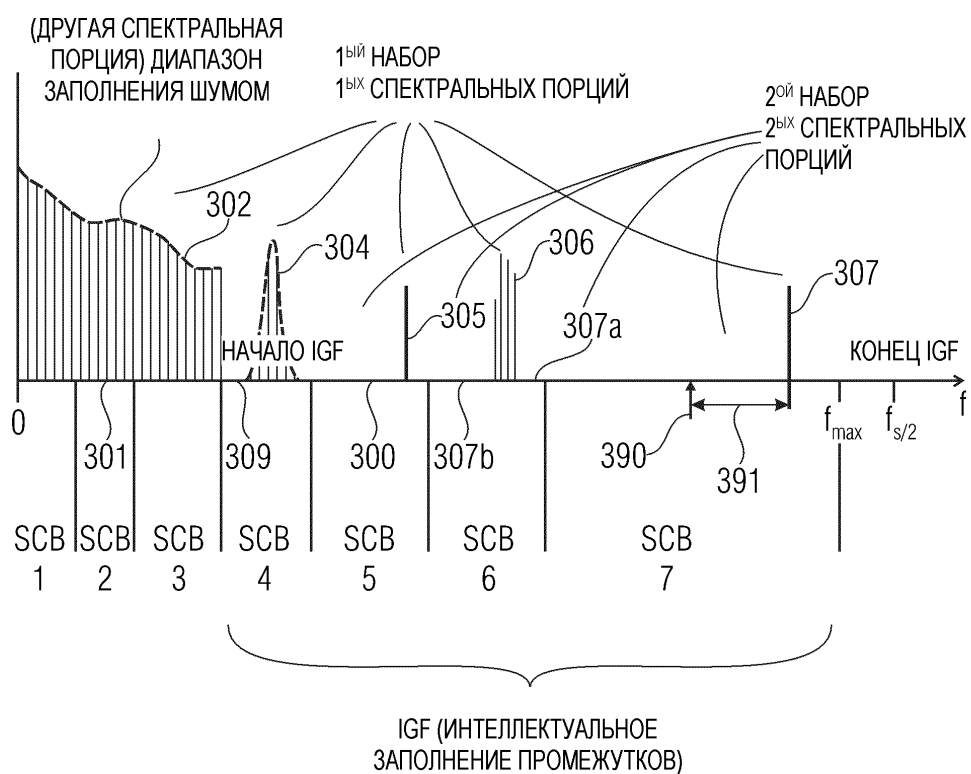
ФИГ. 2А



ФИГ. 2В

4/17

- 1^{ое} РАЗРЕШЕНИЕ (ВЫСОКОЕ РАЗРЕШЕНИЕ) ДЛЯ «ОГИБАЮЩЕЙ» 1^{ого} НАБОРА (ПО-ЛИНЕЙНОЕ КОДИРОВАНИЕ)
- 2^{ое} РАЗРЕШЕНИЕ (НИЗКОЕ РАЗРЕШЕНИЕ) ДЛЯ «ОГИБАЮЩЕЙ» 2^{ого} НАБОРА (МАСШТАБНЫЙ КОЭФИЦИЕНТ НА КАЖДЫЙ SCB)



ФИГ. 3А

5/17

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

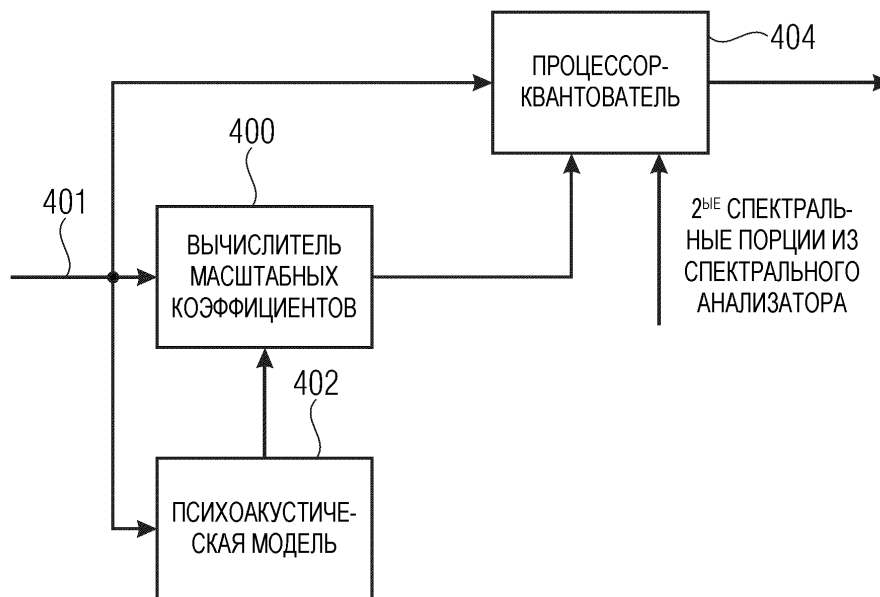
308

310

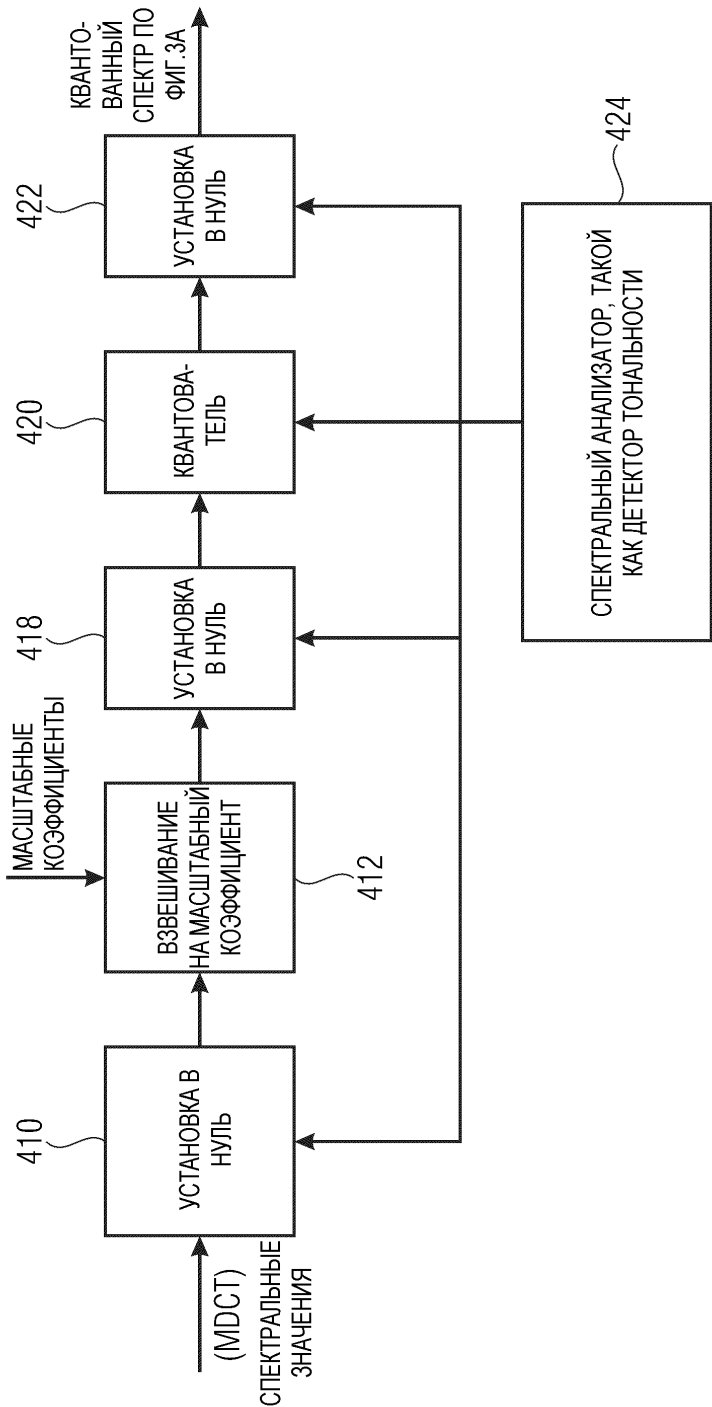
312

ФИГ. 3В

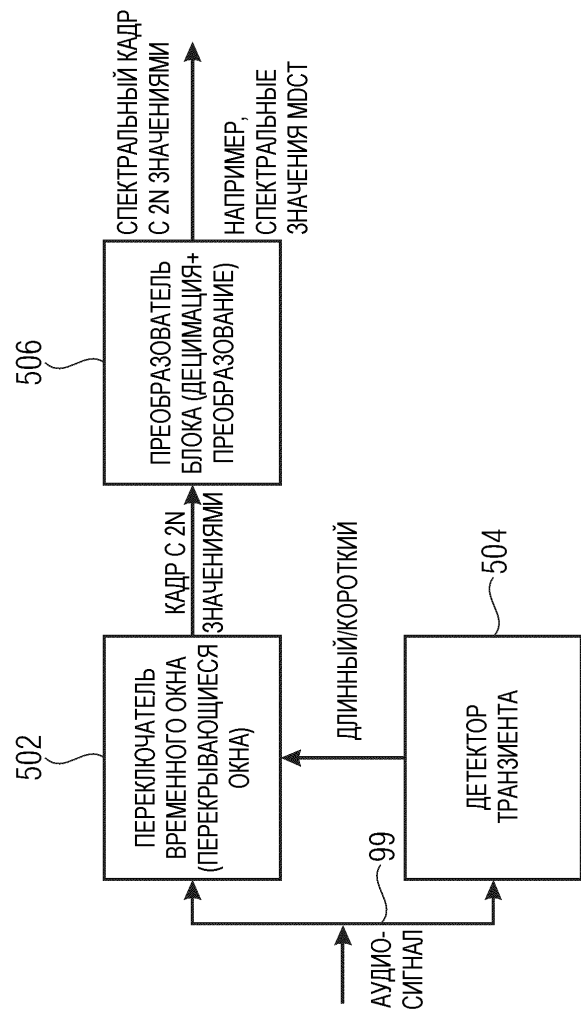
6/17



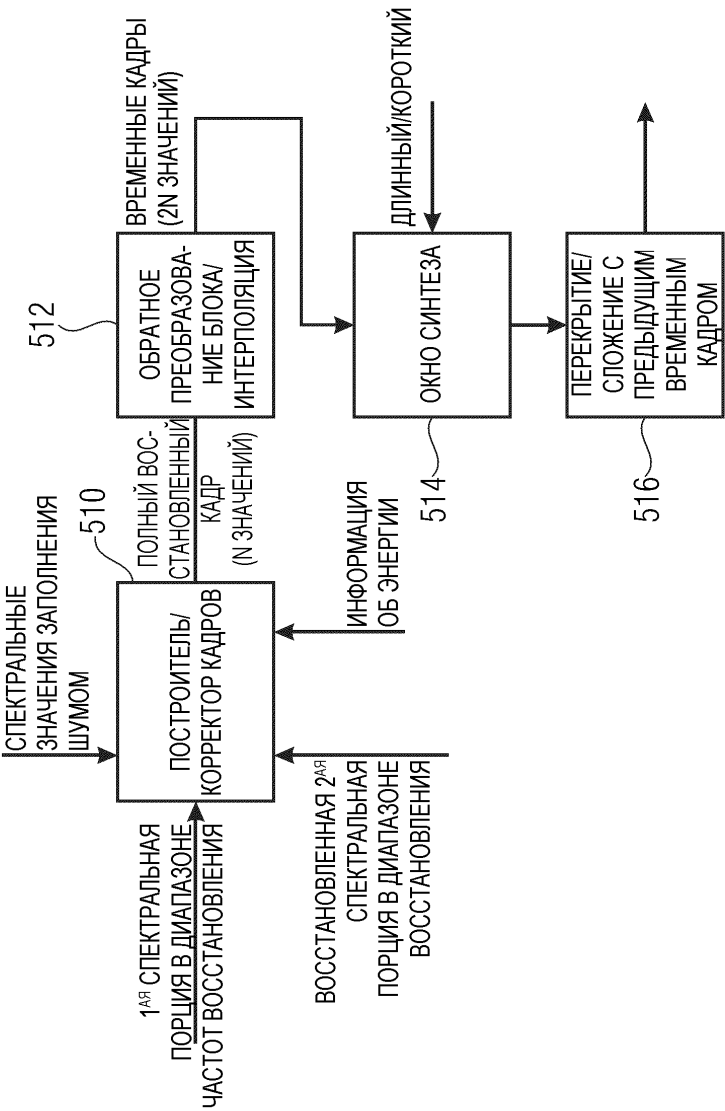
ФИГ. 4А



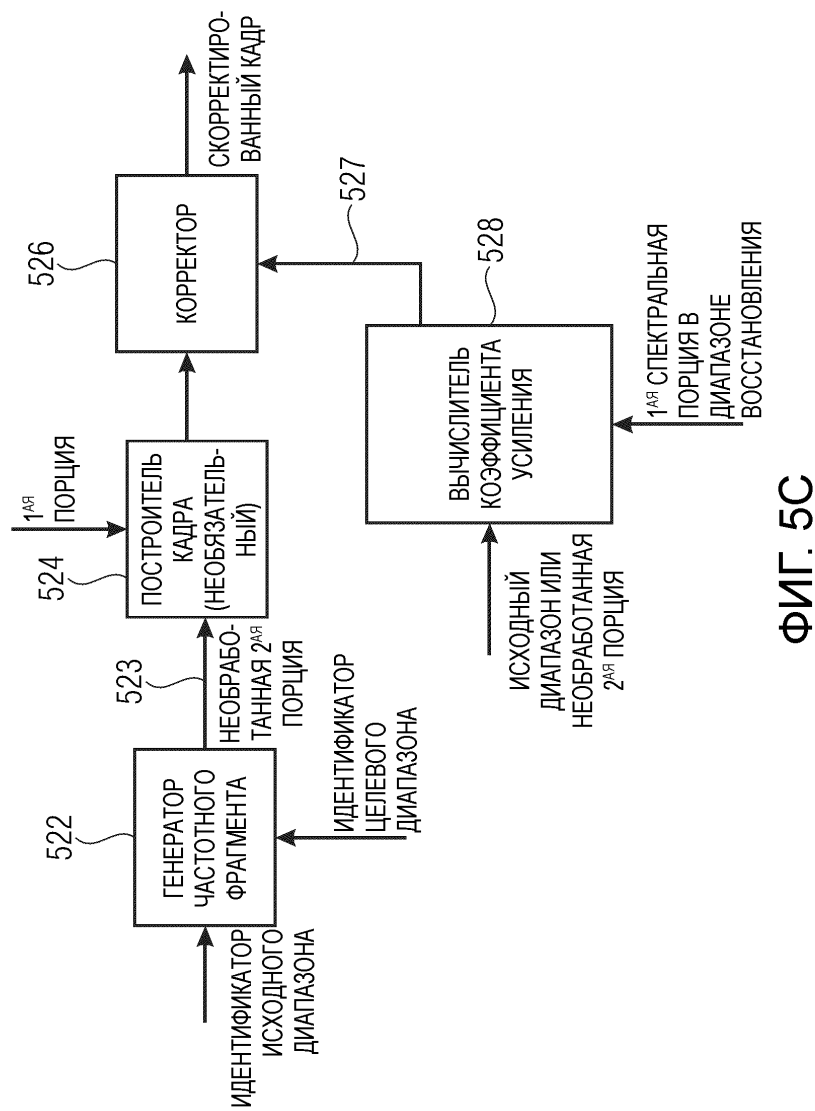
ФИГ. 4В
(ПРОЦЕССОР-КВАНТОВАТЕЛЬ)



ФИГ. 5А
(ДРУГИЕ СПЕКТРАЛЬНЫЕ ПОРЦИИ)

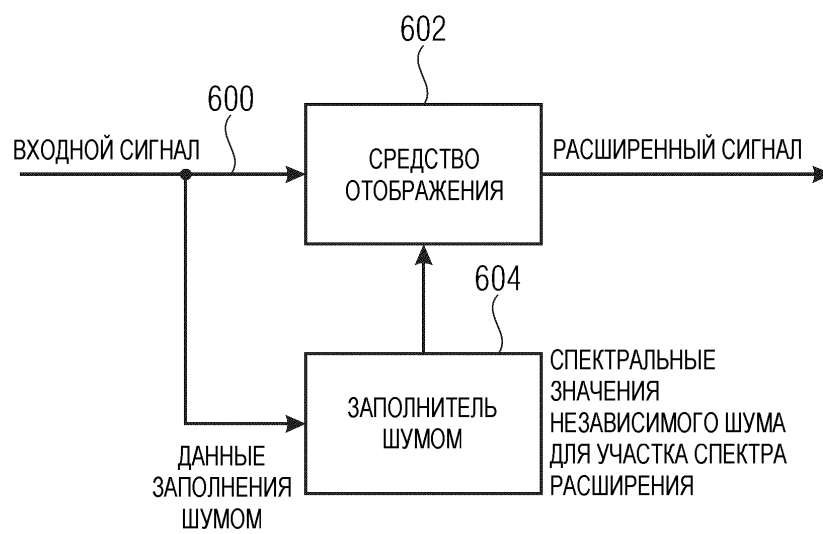


ФИГ. 5В

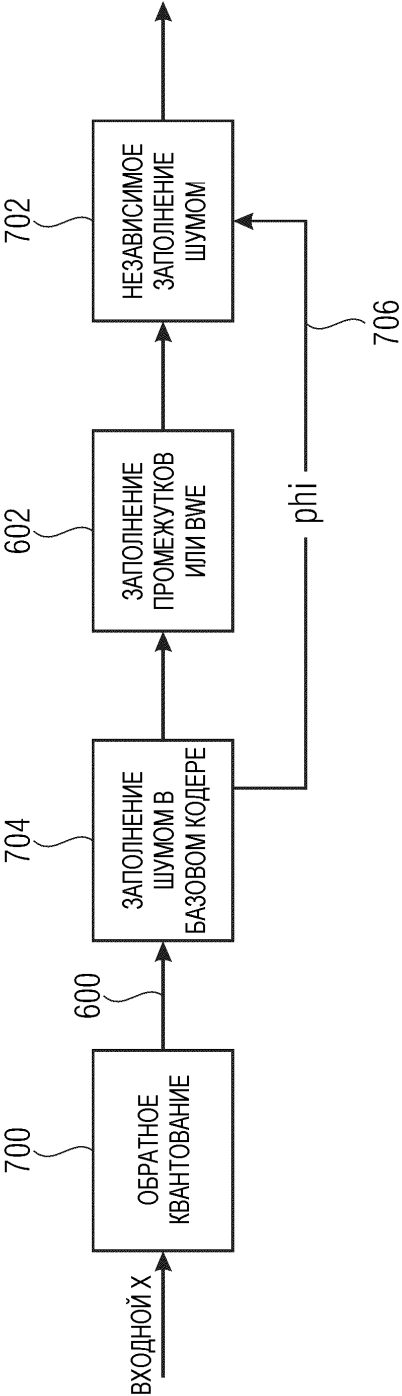


ФИГ. 5С

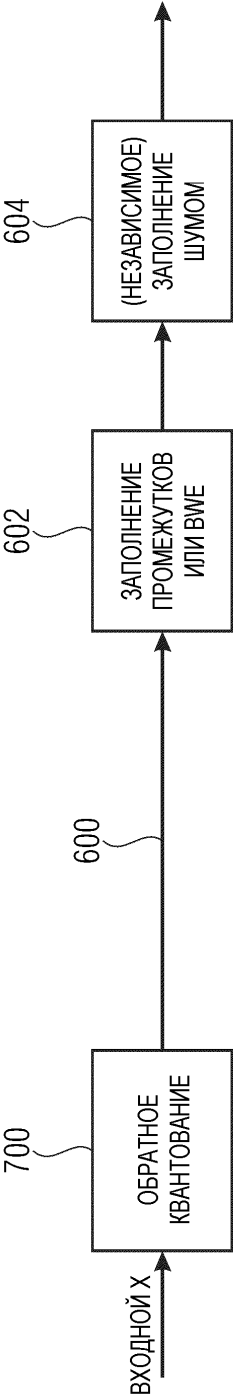
11/17



ФИГ. 6

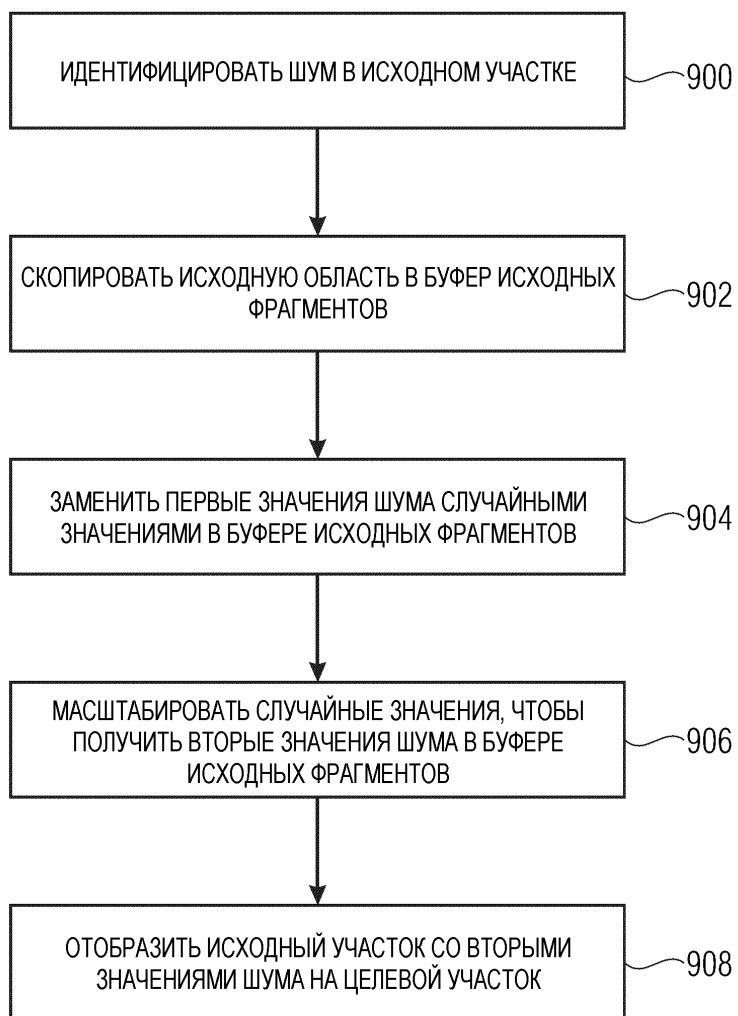


ФИГ. 7



ФИГ. 8

13/17



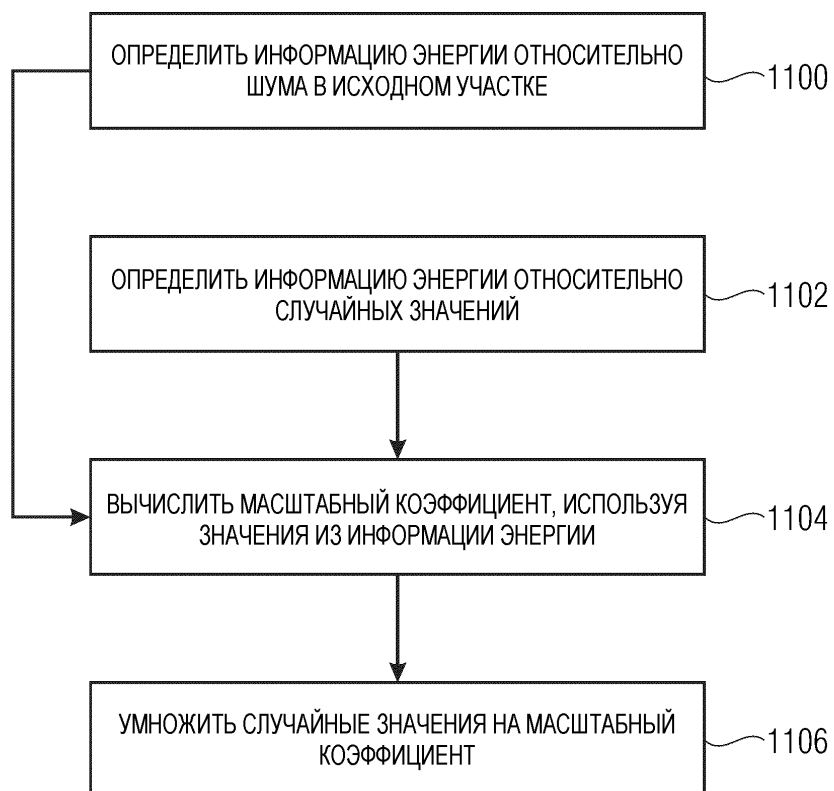
ФИГ. 9

14/17



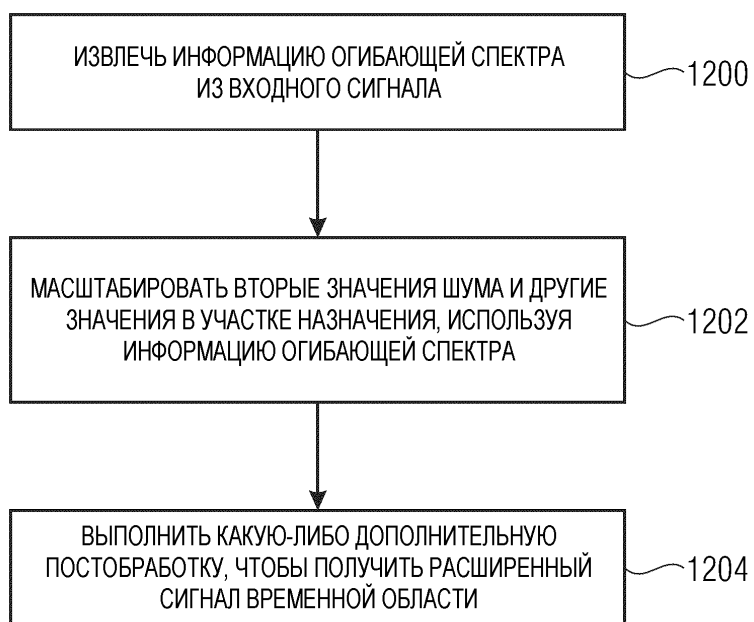
ФИГ. 10

15/17



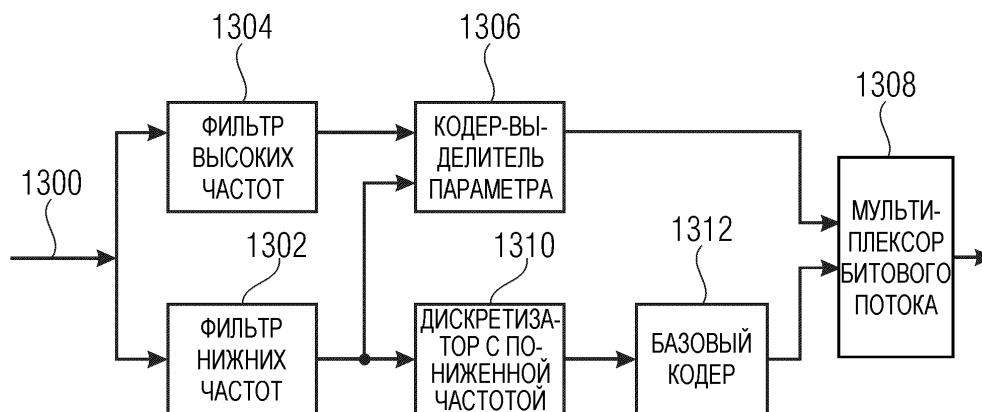
ФИГ. 11

16/17

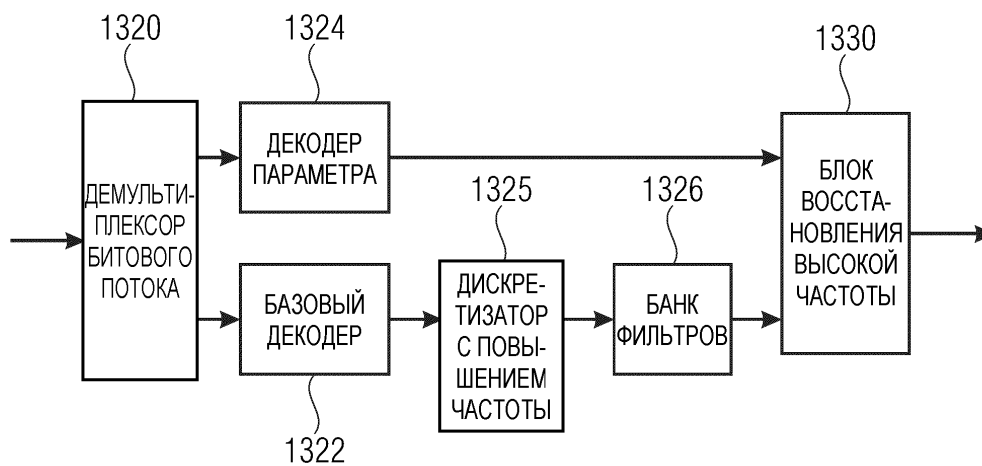


ФИГ. 12

17/17



ФИГ. 13А



ФИГ. 13В