



**ФЕДЕРАЛЬНАЯ СЛУЖБА  
ПО ИНТЕЛЛЕКТУАЛЬНОЙ СОБСТВЕННОСТИ**

**(12) ОПИСАНИЕ ИЗОБРЕТЕНИЯ К ПАТЕНТУ**

(21)(22) Заявка: 2011101617/08, 23.06.2009

(24) Дата начала отсчета срока действия патента:  
23.06.2009

Приоритет(ы):

(30) Конвенционный приоритет:  
11.07.2008 US 61/079,841

(43) Дата публикации заявки: 27.07.2012 Бюл. № 21

(45) Опубликовано: 10.07.2013 Бюл. № 19

(56) Список документов, цитированных в отчете о  
поиске: US 2008/0120116 A1, 22.05.2008. US  
2006/0256971 A1, 16.11.2006. EP 1216474 B1,  
14.07.2004. RU 2256293 C2, 10.07.2005.(85) Дата начала рассмотрения заявки РСТ на  
национальной фазе: 18.01.2011(86) Заявка РСТ:  
EP 2009/004523 (23.06.2009)(87) Публикация заявки РСТ:  
WO 2010/003546 (14.01.2010)Адрес для переписки:  
410000, г. Саратов, главпочтамт, а/я 62,  
ООО "ПатентВолгаСервис", Н.В. Романовой

(72) Автор(ы):

**ГРИЛЛ Бернхард (DE),  
КРАЕМЕР Ульрих (DE),  
МУЛТРУС Маркус (DE),  
НУЕНДОРФ Макс (DE),  
ПОПП Харальд (DE),  
РЕТТЕЛБАХ Николаус (DE),  
НАГЕЛЬ Фредерик (DE),  
ЛОХВАССЕР Маркус (DE),  
ГАЙЕР Марк (DE),  
ЯНДЕР Мануэль (DE),  
БАЧИГАЛУПО Вирджилио (DE)**

(73) Патентообладатель(и):

**Фраунхофер-Гезелльшафт цур Фёрдерунг  
дер ангевандтен Форшунг Е.Ф. (DE)**

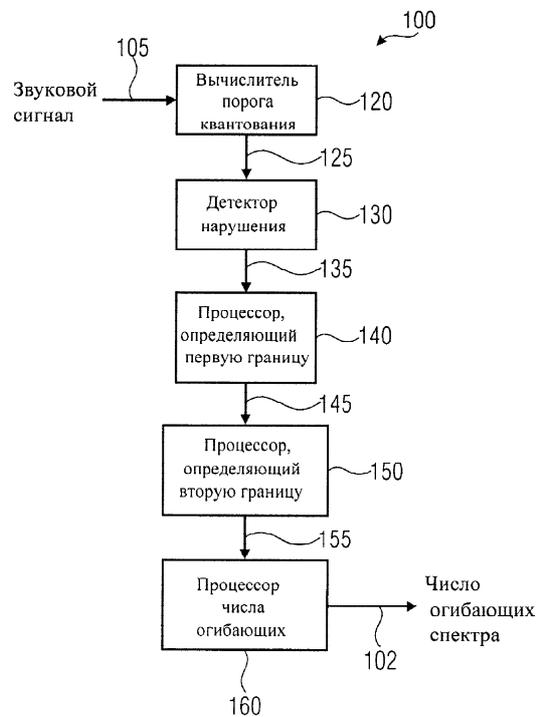
**(54) УСТРОЙСТВО И СПОСОБ ДЛЯ ВЫЧИСЛЕНИЯ ЧИСЛА ОГИБАЮЩИХ СПЕКТРА**

(57) Реферат:

Изобретение относится к области вычисления числа огибающих спектра, а именно к кодированию звуковых сигналов. Техническим результатом является эффективное кодирование в лучшем качестве, специально для сигналов с медленно изменяющейся энергией, интенсивность колебаний которой слишком низкая, чтобы быть обнаруженной обычными детекторами кратковременных помех. Устройство для вычисления числа огибающих спектра включает: вычислитель порога квантования; детектор для обнаружения нарушения пороговой величины при помощи порога

квантования; процессор для определения первой границы огибающей между парой соседних временных частей; процессор для определения второй границы огибающей между отличной парой соседних временных частей; числовой процессор для установления числа огибающих спектра, имеющих первую границу огибающей и вторую границу огибающей; переключающийся блок принятия решения, формируемый, чтобы предоставить сигнал переключения принятия решения; сигнал переключения принятия решения подает речеподобный звуковой сигнал и обычный звукоподобный звуковой сигнал, где детектор приспособлен, чтобы понижать

пороговую величину для речеподобных звуковых сигналов. Способ описывает работу данного устройства. 4 н. и 8 з.п. ф-лы, 11 ил.



Фиг. 1

RU 2 4 8 7 4 2 8 C 2

RU 2 4 8 7 4 2 8 C 2



FEDERAL SERVICE  
FOR INTELLECTUAL PROPERTY

(12) **ABSTRACT OF INVENTION**

(21)(22) Application: **2011101617/08, 23.06.2009**

(24) Effective date for property rights:  
**23.06.2009**

Priority:

(30) Convention priority:  
**11.07.2008 US 61/079,841**

(43) Application published: **27.07.2012 Bull. 21**

(45) Date of publication: **10.07.2013 Bull. 19**

(85) Commencement of national phase: **18.01.2011**

(86) PCT application:  
**EP 2009/004523 (23.06.2009)**

(87) PCT publication:  
**WO 2010/003546 (14.01.2010)**

Mail address:

**410000, g. Saratov, glavpochtamt, a/ja 62, OOO "PatentVolgaServis", N.V. Romanovoj**

(72) Inventor(s):

**GRILL Bernhard (DE),  
KRAEMER Ul'rikh (DE),  
MULTRUS Markus (DE),  
NUENDORF Maks (DE),  
POPP Kharal'd (DE),  
RETTELBAKh Nikolaus (DE),  
NAGEL' Frederik (DE),  
LOKhVASSER Markus (DE),  
GAJER Mark (DE),  
JaNDER Manuehl' (DE),  
BACHIGALUPO Virdzhilio (DE)**

(73) Proprietor(s):

**Fraunkhofer-Gezell'shaft tsur Ferderung der  
angevandten Forshung E.F. (DE)**

(54) **APPARATUS AND METHOD FOR CALCULATING NUMBER OF SPECTRAL ENVELOPES**

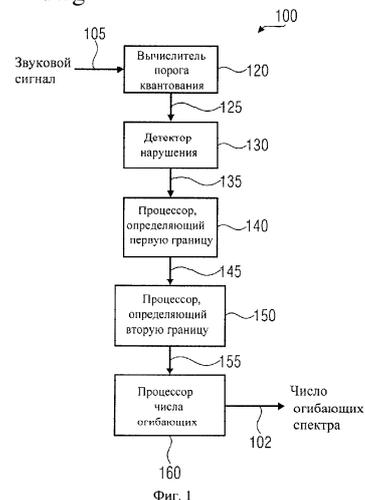
(57) Abstract:

FIELD: information technology.

SUBSTANCE: apparatus for calculating the number of spectral envelopes includes: a quantisation threshold calculator; a detector for detecting violation of the threshold value using the quantisation threshold; a processor for determining a first envelope border between the pair of neighbouring time portions; a processor for determining a second envelope border between a different pair of neighbouring time portions; a number processor for establishing the number of spectral envelopes having the first envelope border and the second envelope border; a switching decision unit formed to provide a decision switching signal; the decision switching signal yields a speech-like audio signal and a common sound-like audio signal, where the detector is capable of lowering the threshold value for speech-like audio signals. The method describes operation of said apparatus.

EFFECT: efficient encoding in the best quality, especially for signals with slow-varying energy, oscillation intensity of which is too to be detected by ordinary short-time interference detectors.

12 cl, 11 dwg



RU 2 487 428 C2

RU 2 487 428 C2

Данное изобретение имеет отношение к устройству и способу вычисления числа огибающих спектра, к звуковому кодирующему устройству и способу кодирования звуковых сигналов.

5 Естественное звуковое кодирование и речевое кодирование - две главные задачи кодер-декодеров для звуковых сигналов. Естественное звуковое кодирование обычно используется для музыкальных или произвольных сигналов при средних скоростях  
10 передачи битов и обычно предлагает широкие звуковые полосы пропускания. С другой стороны, речевые кодирующие устройства в основном ограничиваются воспроизведением речи, но могут также использоваться при очень низкой скорости  
15 передачи битов. Широкополосная речь предлагает главное субъективное повышение качества на узкой полосе речи. Увеличение полосы пропускания не только улучшает ясность и естественность речи, но также и распознавание спикера. Широкополосное речевое кодирование, таким образом, - важная проблема для следующего поколения  
20 телефонных структур. Далее, из-за огромного роста мультимедийной сферы, весьма желательно передавать музыкальные и другие неречевые сигналы по телефонным системам с высоким качеством.

Чтобы радикально уменьшить скорость передачи битов, кодирование источника  
20 может выполняться посредством использования расщепляющих полосу перцепционных звуковых кодер-декодеров. Эти естественные звуковые кодер-декодеры эксплуатируют перцепционную ненужность и статистическую избыточность в сигнале. Кроме того, распространено уменьшение частоты дискретизации и, таким  
25 образом, звуковой полосы пропускания. Также распространено сокращение числа структурных уровней, что иногда обеспечивает слышимое искажение квантизации, и использование деградации стереообласти во время кодирования интенсивности. Злоупотребление такими методами приводит к раздражающей перцепционной  
30 деградации. Чтобы улучшить исполнение кодирования, в качестве эффективного метода используется репликация спектральной полосы, чтобы генерировать высокочастотные сигналы в кодер-декодере, основанном на высокочастотной реконструкции (HFR).

Репликация спектральной полосы (SBR) включает методику, завоевавшую  
35 популярность в качестве дополнения к популярным перцепционным звуковым кодирующим устройствам, таким как MP3 и улучшенное звуковое кодирование (AAC). SBR включает способ расширения полосы пропускания, в котором нижняя полоса (базовая полоса или основная полоса) спектра кодируется посредством использования современных кодер-декодеров, тогда как верхняя полоса (или высокая полоса) грубо  
40 параметризуются посредством использования нескольких параметров. SBR использует корреляцию между нижней полосой и верхней полосой посредством прогнозирования более широкой полосы сигнала от более низкой полосы, используя извлеченные характеристики верхней полосы. Этого часто бывает достаточно, так как  
45 человеческое ухо менее чувствительно к искажениям в более высокой полосе по сравнению с более низкой полосой. Новые звуковые кодирующие устройства, поэтому, кодируют более низкий спектр, используя, например, MP3 или AAC, тогда как верхняя полоса кодируется при помощи SBR. Ключом к алгоритму SBR является информация, используемая для описания более высокой частотной части сигнала.  
50 Главная цель разработки этого алгоритма состоит в том, чтобы восстановить спектр более высокой полосы, не вводя артефактов, и чтобы обеспечить хорошее спектральное и временное разрешение. Например, 64-полосная комплекснозначная многофазная гребенка фильтров используется в анализирующем блоке и в

кодирующем устройстве; гребенка фильтров используется, чтобы получить, например, образцы энергии верхней полосы оригинального входного сигнала. Эти образцы энергии могут тогда использоваться как опорные величины для схемы регулирования огибающей, используемой в декодере.

5 Огибающие спектра обращаются к грубому спектральному распределению сигнала в общем смысле и включают, например, коэффициенты фильтрации в линейном кодирующем устройстве, основанном на прогнозе, или множестве средних по частоте / времени образцов поддиапазона в кодирующем устройстве поддиапазона. Данные  
10 огибающей обращаются, в свою очередь, к квантованной и кодированной огибающей спектра. В особенности, если более низкий частотный диапазон кодируется с низкой скоростью передачи битов, данные огибающей составляют большую часть битового потока. Следовательно, важно сжато представить огибающую спектра, используя особенно низкие скорости передачи битов.

15 Репликация спектральной полосы использует инструменты, которые основываются на репликации, например, последовательности гармоник, усеченных во время кодирования. Кроме того, таким образом регулируется огибающая спектра генерированной верхней полосы и применяется обратное фильтрование и  
20 добавляются шумовые и гармонические компоненты, чтобы восстановить спектральные характеристики оригинального сигнала. Поэтому вход инструмента SBR включает, например, квантованные данные огибающей, различные управляющие данные, сигнал временной области от основного кодирующего устройства (например, AAC или MP3). Выход инструмента SBR - или сигнал временной области, или  
25 представление QMF-области ( $QMF = \text{Квадратурный зеркальный фильтр}$ ) сигнала, как, например, в случае, если используется MPEG окружающий инструмент. Описание элементов битового потока для полезной нагрузки SBR может быть найдено в Стандарте ISO/IEC 14496-3:2005, подпункт 4.5.2.8, оно включает, среди других данных,  
30 данные расширения SBR, заголовок SBR и указывает число огибающих SBR в пределах SBR фрейма.

Для выполнения SBR на стороне кодирующего устройства анализ выполняется на входном сигнале. Информация, полученная из этого анализа, используется, чтобы выбрать подходящее временное / частотное разрешение данного SBR фрейма.  
35 Алгоритм вычисляет границы начального момента времени и конечного момента времени огибающих SBR в данном SBR фрейме, число огибающих SBR, а также их частотное разрешение. Различные частотные разрешения вычисляются, как описано, например, в Стандарте ISO/IEC 144963 в подпункте 4.6.18.3. Алгоритм также  
40 вычисляет число минимальных уровней шума для данного SBR фрейма и его границы начального момента времени и конечного момента времени. Границы начального момента времени и конечного момента времени минимальных уровней шума должны быть подмножеством границ начального момента времени и конечного момента времени огибающих спектра. Алгоритм делит данный SBR фрейм на четыре класса:

45 FIXFIX - и ведущая, и замыкающая временные границы равны номинальным границам SBR-фрейма. Все временные границы огибающей SBR в фрейме однородно распределены во времени. Число огибающих - целочисленная степень двух (1, 2, 4, 8, ...).

50 FIXVAR - ведущая временная граница равняется ведущей номинальной границе фрейма. Замыкающая временная граница является переменной и может быть определена элементами битового потока. Все временные границы огибающей SBR между ведущей и замыкающей временными границами могут быть определены как

относительное расстояние в квантах времени до предыдущей границы, начиная с замыкающей временной границы.

VARFIX - ведущая временная граница является переменной и может определяться элементами битового потока. Замыкающая временная граница равняется замыкающей номинальной границе структуры. Все временные границы огибающей SBR между ведущей и замыкающей временными границами определяются в битовом потоке как относительное расстояние в квантах времени до предыдущей границы, начиная с ведущей временной границы.

VARVAR - и ведущая, и замыкающая временные границы являются переменными и могут определяться в битовом потоке. Определяются также все временные границы огибающей SBR между ведущей и замыкающей временными границами.

Относительные временные границы, начиная с ведущей временной границы, определяются как относительное расстояние до предыдущей временной границы.

Относительные временные границы, начиная с замыкающей временной границы, определяются как относительное расстояние до предыдущей временной границы.

Нет никаких ограничений на переходы классов SBR фрейма, то есть в Стандарте допустима любая последовательность классов. Однако, в соответствии с этим Стандартом, максимальное число огибающих SBR на SBR фрейм ограничено 4 для класса FIXFIX и 5 для класса VARVAR. Классы FIXVAR и VARFIX синтаксически ограничены четырьмя огибающими SBR. Огибающие спектра SBR фрейма оцениваются на временном сегменте и с частотным разрешением, предоставляемыми временной/частотной сеткой. Огибающая SBR оценивается посредством усреднения возведенных в квадрат сложных образцов поддиапазона в данных временных/частотных областях.

Обычно, кратковременные помехи получают в SBR определенную обработку посредством использования определенных огибающих переменных длин.

Кратковременные помехи могут определяться частями в пределах обычных сигналов, где сильное увеличение энергии появляется в пределах короткого промежутка времени, который может быть или может не быть ограничен в определенной частотной области. Примеры кратковременных помех - удары кастаньет и ударных инструментов, а также определенные звуки человеческого голоса, как, например, буквы: П, Т, К. Обнаружение такого рода кратковременных помех пока всегда осуществляется таким образом или посредством того же самого алгоритма (использующего переходную пороговую величину), который независим от сигнала, если он классифицируется как речь, или классифицируется как музыка. Кроме того, возможное различие между вокализованной и невокализованной речью не влияет на обычный или классический механизм обнаружения кратковременных помех.

Следовательно, в случае обнаружения кратковременной помехи, SBR-данные должны быть приспособлены так, чтобы декодер мог должным образом копировать обнаруженную кратковременную помеху. В WO 01/26095 устройство и способ раскрываются для кодирования огибающей спектра, которое принимает во внимание обнаруженную кратковременную помеху в звуковом сигнале. В этом традиционном способе неоднородная временная и частотная выборка огибающей спектра достигается посредством адаптивно группирующихся образцов поддиапазона из гребенки фильтров установленного размера в частотных диапазонах и временных сегментах, каждый из которых генерирует один образец огибающей.

Соответствующая система устанавливается по умолчанию на долговременные сегменты и высокочастотное разрешение, но около кратковременной помехи

используются более короткие временные сегменты, посредством чего могут использоваться большие перепады частот, чтобы поддерживать размер данных в определенных рамках. В случае, если обнаруживается кратковременная помеха, система переключается с FIXFIX фрейма на FIXVAR фрейм, за которым  
5 следует VARFIX фрейм таким образом, что граница огибающей устанавливается прямо перед обнаруженной кратковременной помехой. Эта процедура повторяется всякий раз, когда обнаруживается кратковременная помеха.

В случае, если колебание энергии изменяется только медленно, детектор  
10 кратковременных помех не обнаружит изменение. Эти изменения могут, однако, быть достаточно сильными, чтобы генерировать заметные артефакты, если не будут обработаны должным образом. Простым решением было бы понижение пороговой величины в детекторе кратковременных помех. Это, однако, привело бы к частому  
15 переключению между различными фреймами (FIXFIX на FIXVAR+VARFIX). Как следствие, значительное количество дополнительных данных должно быть передано, подразумеваемая плохую эффективность кодирования, - особенно, если медленное увеличение происходит за более длительное время (например, на множественных фреймах). Это не приемлемо, так как сигнал не имеет такую сложность, которая  
20 оправдала бы более высокую скорость передачи данных, и, следовательно, эта опция не решит проблему.

Задачей данного изобретения, поэтому, является предоставление устройства, которое обеспечит эффективное кодирование без заметных артефактов, особенно для сигналов, включающих медленно изменяющуюся энергию, которая слишком низка,  
25 чтобы быть обнаруженной детектором кратковременных помех.

Эта задача достигается устройством по п.1, кодирующим устройством по п.11, способом вычисления числа огибающих спектра по п.13 или способом генерирования потока данных по п.14.

30 Данное изобретение основывается на обнаружении того, что перцепционное качество переданного звукового сигнала может быть улучшено посредством гибкого регулирования числа огибающих спектра в пределах SBR фрейма в соответствии с данным сигналом. Это достигается сравнением звукового сигнала соседних временных частей в пределах SBR фрейма. Сравнение осуществляется посредством  
35 определения распределения энергии для звукового сигнала в пределах временных частей, и порог квантования измеряет отклонение распределений энергии двух соседних временных частей. В зависимости от того, прерывает ли порог квантования пороговую величину, граница огибающей располагается между соседними  
40 временными частями. Другая граница огибающей может быть или вначале, или в конце SBR фрейма или, альтернативно, также между двумя дальнейшими соседними временными частями в пределах SBR фрейма.

В результате, SBR фрейм не приспособливается или не изменяется, как, например, в обычном устройстве, где изменение FIXFIX фрейма на FIXVAR фрейм или на VARFIX  
45 фрейм выполняется, чтобы обработать кратковременные помехи. Вместо этого осуществления используют переменное число огибающих, например, в пределах FIXFIX фреймов, чтобы принять во внимание переменные колебания звукового сигнала так, чтобы даже медленно изменяющиеся сигналы могли привести  
50 к изменению числа огибающих и, к тому же, обеспечить улучшенное звуковое качество, производимое инструментом SBR в декодере. Определенные огибающие могут, например, покрывать части равной продолжительности в пределах SBR фрейма. Например, SBR фрейм может быть разделен на предварительно определенное

число временных частей (которое может, например, включить 4, 8 или другую целочисленную степень числа 2).

Распределение спектральной энергии каждой временной части может покрывать только верхний частотный диапазон, который копируется SBR. С другой стороны, распределение спектральной энергии может также быть связано с целым частотным диапазоном (верхним и нижним), где верхний частотный диапазон может быть или не быть взвешен больше, чем нижний частотный диапазон. В соответствии с этой процедурой уже одного нарушения пороговой величины может быть достаточно, чтобы увеличить число огибающих, или чтобы использовать максимальное число огибающих в пределах SBR фрейма.

Дальнейшие осуществления могут также включать инструмент классификатора сигнала, который анализирует оригинальный входной сигнал и генерирует из него управляющую информацию, которая запускает выбор различных кодирующих режимов. Различные кодирующие режимы могут, например, включать речевое кодирующее устройство и обычное звуковое кодирующее устройство. Анализ входного сигнала зависит от исполнения с целью выбрать оптимальный основной кодирующий режим для данного фрейма входного сигнала. Оптимальный вариант связан с балансированием перцепционного высокого качества при использовании только низкой скорости передачи битов для кодирования. Входом инструмента классификатора сигнала может быть оригинальный неизменный входной сигнал и/или дополнительные зависимые от исполнения параметры. Выходом инструмента классификатора сигнала может, например, быть управляющий сигнал для управления выбором основного кодер-декодера.

Если, например, сигнал идентифицирован или классифицирован как речь, времениподобное разрешение расширения полосы пропускания (BWE) может быть увеличено (например, большим количеством огибающих) так, чтобы времениподобное колебание энергии (медленно или сильно колеблющееся) могло лучше приниматься во внимание.

Этот подход принимает во внимание то, что различные сигналы с различными временными/частотными характеристиками имеют различные требования относительно характеристик расширения полосы пропускания. Например, сигналы кратковременных помех (появляющиеся, например, в речевых сигналах) нуждаются в высоком временном разрешении BWE, частота деления (которая означает верхнюю частотную границу основного кодирующего устройства) должна быть насколько возможно высокой. Особенно в вокализованной речи искаженная временная структура может снизить качество восприятия. С другой стороны, тональные сигналы часто нуждаются в устойчивом воспроизведении спектральных компонентов и согласованного гармонического рисунка воспроизведенных высокочастотных частей. Устойчивое воспроизведение тональных частей ограничивает полосу пропускания основного кодирующего устройства - оно не нуждается в BWE с высоким временным разрешением, но вместо этого с более высоким спектральным разрешением. В проекте, обеспечивающем переключение основного кодирующего устройства с речи на звук, кроме того, возможно использовать решение основного кодирующего устройства, чтобы приспособить как временные, так и спектральные характеристики BWE, а также, чтобы приспособить полосу пропускания основного кодирующего устройства к характеристикам сигнала.

Если все огибающие включают ту же самую продолжительность, зависящую от обнаруженного нарушения (в какой-то момент времени), число огибающих может

отличаться от фрейма к фрейму. Осуществления определяют число огибающих для SBR фрейма, например, следующим образом. Можно начать с разделения максимально возможного числа огибающих (например, 8) и постепенно сократить число огибающих так, чтобы в зависимости от входного сигнала не использовалось больше огибающих, чем необходимо, чтобы обеспечить воспроизведение сигнала в перцепционно высоком качестве.

Например, нарушение, обнаруженное уже на первой границе временных частей в пределах фрейма, может дать в результате максимальное число огибающих, тогда как нарушение, обнаруженное только на второй границе, может дать в результате половину максимального числа огибающих. Чтобы уменьшить количество данных, подлежащих передаче, в дальнейших осуществлениях пороговая величина может зависеть от момента времени (то есть в зависимости от того, какая граница в настоящее время анализируется). Например, между первой и второй временными частями (первая граница) и между третьей и четвертой временными части (третья граница) пороговая величина в обоих случаях может быть выше, чем между второй и третьей временными частями (вторая граница). Таким образом, статистически будет больше нарушений на второй границе, чем на первой или третьей границе, и, следовательно, более вероятно, меньше огибающих, что было бы предпочтительнее (более детально см. ниже).

В дальнейших осуществлениях продолжительность временной части предварительно определенного числа последующих временных частей равна минимальной продолжительности, для которой определяется единственная огибающая, и в которой вычислитель порога квантования приспособливается, чтобы вычислять порог квантования для двух соседних временных частей, имеющих минимальную продолжительность.

Дальнейшие осуществления включают информационный процессор для предоставления дополнительной информации; дополнительная информация включает первую границу огибающей и вторую границу огибающей в пределах временной последовательности звукового сигнала. В дальнейших осуществлениях детектор приспособливается, чтобы исследовать во временной последовательности каждую границу между соседними временными частями.

Осуществления также используют устройство для вычисления числа огибающих в пределах кодирующего устройства. Кодирующее устройство включает устройство для вычисления числа огибающих спектра, а вычислитель огибающих использует это число, чтобы вычислять данные огибающей спектра для SBR фрейма. Осуществления также включают способ вычисления числа огибающих и способ кодирования звукового сигнала.

Поэтому использование огибающих в пределах FIXFIX фрейма направлено на обеспечение улучшенного моделирования колебания энергии, которое не покрывается указанными обработками кратковременных помех, так как они слишком медленные, чтобы быть обнаруженными как кратковременные помехи или быть классифицированными как кратковременные помехи. С другой стороны, они достаточно быстрые, чтобы вызвать появление артефактов, если их не обрабатывать должным образом, из-за недостаточного времени подобного разрешения. Поэтому обработка огибающих, согласно данному изобретению, принимает во внимание медленно изменяющиеся колебания энергии, а не только сильные или быстрые колебания энергии, которые характерны для кратковременных помех. Следовательно, осуществления данного изобретения обеспечивают более эффективное кодирование в

лучшем качестве, специально для сигналов с медленно изменяющейся энергией, интенсивность колебаний которой слишком низкая, чтобы быть обнаруженной обычными детекторами кратковременных помех.

Краткое описание чертежей

5 Данное изобретение будет теперь описано и проиллюстрировано примерами. Характерные черты изобретения будут легко оценены и лучше поняты со ссылкой на следующее детальное описание, которое следует рассматривать со ссылкой на сопровождающие чертежи, на которых:

10 Фиг.1 показывает блок-схему устройства для вычисления числа огибающих спектра согласно осуществлению данного изобретения;

Фиг.2 показывает блок-схему модуля SBR, включающего вычислитель числа огибающих;

15 Фиг.3А и 3В показывают блок-схемы кодирующего устройства, включающего вычислитель числа огибающих;

Фиг.4 иллюстрирует разделение SBR фрейма на предварительно определенное число временных частей;

20 Фиг.5а-5с показывают дальнейшее разделение SBR фрейма, включающего три огибающих, покрывающих различные числа временных частей;

Фиг.6А и 6В иллюстрируют распределение спектральной энергии для сигналов в пределах соседних временных частей; и

25 Фиг.7А-7С показывают кодирующее устройство, включающее дополнительное переключение звук/речь, вызывающее различное временное разрешение звукового сигнала.

Детальное описание изобретения

30 Осуществления, описанные ниже, просто иллюстрируют принцип данного изобретения для улучшения репликации спектральной полосы, например, используемой в звуковом кодирующем устройстве. Подразумевается, что модификации и изменения расположения и деталей, описанных здесь, будут очевидны для специалистов, квалифицированных в этой области. Поэтому цель состоит в том, чтобы не ограничиваться конкретными деталями, представленными здесь посредством описания и объяснения осуществлений.

35 Фиг.1 показывает устройство 100 для вычисления числа 102 огибающих спектра 104. Огибающие спектра 104 получены посредством кодирующего устройства для репликации спектральной полосы, где кодирующее устройство приспособлено, чтобы кодировать звуковой сигнал 105, используя множество выборочных значений в 40 пределах предварительно определенного числа последующих временных частей 110 в фрейме репликации спектральной полосы (SBR фрейм), простирающийся от начального момента времени  $t_0$  до конечного момента времени  $t_n$ . Предварительно определенное число последующих временных частей 110 расположено во временной последовательности, обусловленной звуковым сигналом 105.

45 Устройство 100 включает вычислитель порога квантования 120 для определения порога квантования 125, где порог квантования 125 измеряет отклонение в распределениях спектральной энергии пары соседних временных частей. Устройство 100 далее включает детектор нарушения 130 для обнаружения 50 нарушения 135 пороговой величины посредством порога квантования 125. Кроме того, устройство 100 включает процессор 140 (процессор, определяющий первую границу) для определения первой границы огибающей 145 между парой соседних временных частей, когда обнаружено нарушение 135 пороговой величины.

Устройство 100 также включает процессор 150 (процессор, определяющий вторую границу) для определения второй границы огибающей 155 между другой парой соседних временных частей или в начальный момент времени  $t_0$  или конечный момент времени  $t_n$  для огибающей 104, имеющей первую границу огибающей 145, основанную на нарушении 135 пороговой величины для другой пары, или основанную на временном положении пары или другой пары в SBR фрейме. Наконец, устройство 100 включает процессор 160 (процессор числа огибающих) для определения числа 102 огибающих спектра 104, имеющих первую границу огибающей 145 и вторую границу огибающей 155.

Дальнейшие осуществления включают устройство 100, в котором продолжительность временной части предварительно определенного числа последующей временной части 110 равна минимальной продолжительности, для которой определяется единственная огибающая 104. Кроме того, вычислитель порога квантования 120 приспособлен для вычисления порога квантования 125 для двух соседних временных частей, имеющих минимальную продолжительность.

Фиг.2 показывает осуществление для инструмента SBR, включающего вычислитель числа огибающих 100 (показанный на фиг.1), который определяет число 102 огибающих спектра 104 посредством обработки звукового сигнала 105. Число 102 вводится в вычислитель огибающих 210, который вычисляет данные огибающей 205 звукового сигнала 105. Используя число 102, вычислитель огибающих 210 разделит SBR фрейм на части, покрываемые огибающей спектра 104, и для каждой огибающей спектра 104 вычислитель огибающих 210 вычисляет данные огибающей 205. Данные огибающей включают, например, квантованную и кодированную огибающие спектра, и эти данные необходимы на стороне декодера для генерирования сигнала высокой полосы и применения обратного фильтрования, добавления шумовых и гармонических компонентов для репликации спектральных характеристик оригинального сигнала.

Фиг.3А показывает осуществление кодирующего устройства 300; кодирующее устройство 300 включает связанные с SBR модули 310, анализирующий блок QMF 320, субдискретизатор 330, основное кодирующее устройство AAC 340 и форматер полезной нагрузки битового потока 350. Кроме того, кодирующее устройство 300 включает вычислитель данных огибающей 210. Кодирующее устройство 300 включает вход для образцов РСМ (звуковой сигнал 105; РСМ=кодово-импульсная модуляция), который соединен с анализирующим блоком QMF 320, и с SBR-связанными модулями 310, и с субдискретизатор 330. Анализирующий блок QMF 320, в свою очередь, соединен с вычислителем данных огибающей 210, который, в свою очередь, соединен с форматером полезной нагрузки битового потока 350.

Субдискретизатор 330 соединен с основным кодирующим устройством AAC 340, который, в свою очередь, соединен с форматером полезной нагрузки битового потока 350. Наконец, SBR-связанный модуль 310 соединен с вычислителем данных огибающей 210 и с основным кодирующим устройством AAC 340.

Поэтому кодирующее устройство 300 субдискретизирует звуковой сигнал 105, чтобы генерировать компоненты в основном частотном диапазоне (в субдискретизаторе 330), которые вводятся в основное кодирующее устройство AAC 340, которое кодирует звуковой сигнал в основном диапазоне частот и передает закодированный сигнал к форматеру полезной нагрузки битового потока 350, в котором кодируемый звуковой сигнал основного частотного диапазона добавляется к закодированному звуковому потоку 355. С другой стороны, звуковой сигнал 105

анализируется анализирующим блоком QMF 320, который извлекает частотные компоненты высокочастотного диапазона и вводит эти сигналы в вычислитель данных огибающей 210. Например, блок QMF с 64 поддиапазонами 320 выполняют поддиапазоновое фильтрование входного сигнала. Выход из гребенки фильтров (то  
 5 есть образцы поддиапазона) является комплекснозначным и, таким образом, супердискретизированным множителем, равным двум, по сравнению со стандартным блоком QMF.

SBR-связанные модули 310 управляют вычислителем данных огибающей 210 посредством предоставления, например, числа 102 огибающих 104 вычислителю данных огибающей 210. Используя число 102 и звуковые компоненты, генерированные анализирующим блоком QMF 320, вычислитель данных огибающей 210 вычисляет данные огибающей 205 и передает данные огибающей 205  
 10 формату полезной нагрузки битового потока 350, который объединяет данные огибающей 205 с компонентами, закодированными основным кодирующим устройством 340, в закодированном звуковом потоке 355.

Фиг.3А показывает, поэтому, часть кодирующего устройства инструмента SBR, оценивающего несколько параметров, используемых в высокочастотном  
 20 восстановлении в декодере.

Фиг.3В показывает пример SBR-связанного модуля 310, который включает вычислитель числа огибающих 100 (показанный на фиг.1) и дополнительно другие модули SBR 360. SBR-связанные модули 310 получают звуковой сигнал 105 и производят число 102 огибающих 104, но также и другие данные, генерированные  
 25 другими модулями SBR 360.

Другие модули SBR 360 могут, например, включать обычный детектор кратковременных помех, приспособленный для обнаружения кратковременных помех в звуковом сигнале 105, и могут также получать число и/или положения огибающих  
 30 так, чтобы модули SBR могли или не могли вычислить часть параметров, используемых способом высокочастотного восстановления в декодере (параметр SBR).

Как было сказано выше, в пределах SBR единица времени SBR (SBR фрейм) может быть разделена на различные блоки данных, так называемые огибающие. Если это подразделение или разделение однородно, то есть если все огибающие 104 имеют тот  
 35 же самый размер, и первая огибающая начинается, а последняя огибающая заканчивается границей структуры, SBR фрейм определяется как FIXFIX фрейм.

Фиг.4 иллюстрирует такое разделение для SBR фрейма на число 102 огибающих спектра 104. SBR фрейм покрывает промежуток времени между начальным моментом  
 40 времени  $t_0$  и конечным моментом времени  $t_n$  и, в осуществлении, показанном на фиг.4, разделяется на 8 временных частей: первая временная часть 111, вторая временная часть 112, ..., седьмая временная часть 117 и восьмая временная часть 118. Восемь временных частей 110 разделены 7-ю границами; это означает, что граница 1 является промежуточной между первой и второй временными частями 111, 112, граница 2  
 45 расположена между второй частью 112 и третьей частью 113, и так далее до границы 7, являющейся промежуточной между седьмой частью 117 и восьмой частью 118.

В Стандарте ISO/IEC 14496-3 максимальное число огибающих 104 в FIXFIX фрейме ограничено четырьмя (см. подраздел 4, параграф 4.6.18.3.6). В общем, число  
 50 огибающих 104 в FIXFIX фрейме может быть степенью двух (например, 1, 2, 4), где FIXFIX фреймы используются только, если в том же самом фрейме не были обнаружены никакие кратковременные помехи. В традиционных выполнениях

высокоэффективных кодирующих устройств ААС, с другой стороны, максимальное число огибающих 104 ограничено двумя, даже если спецификация стандарта теоретически позволяет иметь до четырех огибающих. Это число огибающих 104 на фрейм может быть увеличено, например, до восьми (см. фиг.4), так, чтобы FIXFIX фрейм мог включать 1, 2, 4 или 8 огибающих (или другую степень 2). Конечно, возможно любое другое такое число 102 огибающих 104, чтобы максимальное число огибающих 104 (предварительно определенное число) могло быть ограничено только временным разрешением гребенки фильтров QMF, которая имеет 32 кванта времени QMF на SBR фрейм.

Число 102 огибающих 104 может, например, вычисляться следующим образом. Вычислитель порога квантования 120 измеряет отклонения в распределениях спектральной энергии пар соседних временных частей 110. Например, это означает, что вычислитель порога квантования 120 вычисляет распределение первой спектральной энергии для первой временной части 111, вычисляет распределение второй спектральной энергии из спектральных данных в пределах второй временной части 112, и так далее. Затем, сравнивается распределение первой спектральной энергии и распределение второй спектральной энергии, и из этого сравнения получается порог квантования 125, где порог квантования 125 относится, в этом примере, к границе 1 между первой временной частью 111 и второй временной частью 112. Та же самая процедура может быть применена к второй временной части 112 и к третьей временной части 113 так, чтобы для этих двух соседних временных частей также были получены два распределения спектральной энергии, и эти два распределения спектральной энергии, в свою очередь, сравниваются вычислителем порога квантования 120 для получения дальнейшего порога квантования 125.

В качестве следующего шага, детектор 130 будет сравнивать полученные пороги квантования 125 с пороговой величиной, и если пороговая величина нарушается, то детектор 130 обнаружит нарушение 135. Если детектор 130 обнаруживает нарушение 135, процессор 140 определяет первую границу огибающей 145. Например, если детектор 130 обнаруживает нарушение на границе 1 между первой временной частью 111 и второй временной частью 112, первая граница огибающей 145а располагается на протяжении границы 1.

На фиг.4 осуществление, в котором допустимы только несколько возможностей для гранул/границ, это означало бы, что процесс закончен полностью, и все границы установлены, как обозначено маленькими огибающими, обозначенными цифрами 104а, 104b. В этом случае границы были бы на всех временных моментах 0, 1, 2, ..., n.

Когда, однако, первая граница должна быть установлена, например, на момент времени 4, тогда должен быть произведен поиск второй границы. Как обозначено на фиг.4, вторая граница могла быть на 3, 2, 0. В случае, когда граница находится на 3, процедура закончена полностью, так как установлены наименьшие огибающие 104а, 104b. В случае, когда граница находится на 2, поиск должен быть продолжен, так как еще нет уверенности в том, что могут быть использованы средние огибающие (обозначены цифрой 145а). Даже в случае нахождения границы на 0 еще не определено, что во второй половине, то есть между 4 и n, нет границы. Если во второй половине нет границы, тогда могут быть установлены самые широкие огибающие. Если есть граница, например, на 5, тогда должны использоваться наименьшие огибающие. Если есть граница только на 6, то используются средние огибающие.

Однако, когда допускается полностью гибкий или более гибкий рисунок для огибающих, процедура продолжается, когда была определена первая граница на 1. Тогда процессор 150 определяет вторую границу огибающей 155, которая находится или между другой парой соседних временных частей, или совпадает с начальным моментом времени  $t_0$  или конечным моментом времени  $t_n$ . В осуществлениях, как показано на фиг.4, вторая граница огибающей 155a совпадает с начальным моментом времени  $t_0$  (давая в результате первую огибающую 104a), и другая вторая граница огибающей 155b совпадает с границей 2 между второй временной частью 112 и третьей временной частью 113 (давая в результате вторую огибающую 104b). Если не обнаружено никакое нарушение на границе 1 между первой временной частью 111 и второй временной частью 112, то детектор 130 продолжит исследовать границу 2 между второй временной частью 112 и третьей временной частью 113. Если есть нарушение, другая огибающая 104c простирается от начального момента  $t_0$  до границы 2.

Согласно осуществлениям изобретения для пары соседних огибающих указанный порог квантования 125 измеряет отклонение распределений спектральной энергии, где каждое распределение спектральной энергии относится к части звукового сигнала в пределах временной части. В примере с 8-ю огибающими имеется в сумме 7 измерений (=7 границ между соседними временными частями) или, в общем, если имеется  $n$  огибающих, то имеется  $n-1$  измерений (пороги квантования 125). Каждый из этих порогов квантования 125 может тогда сравниваться с пороговой величиной, и если порог квантования 125 (мера) превышает пороговую величину, то граница огибающей будет расположена между двумя соседними огибающими. В зависимости от определения порога квантования 125 и пороговой величины нарушение может состоять в том, что порог квантования 125 либо выше, либо ниже пороговой величины. В случае, если порог квантования 125 ниже пороговой величины, спектральное распределение может не сильно изменяться от огибающей к огибающей. Следовательно, в этом положении (=момент времени) может не потребоваться никакая граница огибающей.

В предпочтительном осуществлении число 102 огибающих 104 включает степень двух и, кроме того, каждая огибающая включает равный промежуток времени. Это означает, что имеется четыре возможности: первая возможность состоит в том, что целый SBR фрейм покрывается единственной огибающей (не показана на фиг.4), вторая возможность состоит в том, что SBR фрейм покрывается 2 огибающими, третья возможность состоит в том, что SBR фрейм покрывается 4 огибающими, и последняя возможность состоит в том, что SBR фрейм покрывается 8 огибающими (показаны на фиг.4 от основания до вершины).

Может быть полезно исследовать границы в определенном порядке, потому что если имеется нарушение на нечетной границе (граница 1, граница 3, граница 5, граница 7), то число огибающих всегда будет восемь (при условии, что огибающие имеют одинаковый размер). С другой стороны, если имеется нарушение на границе 2 и на границе 6, имеется четыре огибающих, и, наконец, если имеется нарушение только на границе 4, то две огибающие будут закодированы, и если нет никакого нарушения ни на одной из этих 7 границ, весь SBR фрейм покрывается одной единственной огибающей. Следовательно, устройство 100 может исследовать сначала границы 1, 3, 5, 7, и если нарушение обнаруживается на одной из этих границ, устройство 100 может исследовать следующий SBR фрейм, так как в этом случае целый SBR фрейм будет закодирован максимальным числом огибающих. После исследования этих нечетных

5 границ и, если никакие нарушения не обнаружены на нечетных границах, детектор 130 может исследовать, в качестве следующего шага, границу 2 и границу 6 так, что, если нарушение обнаружено на одной из этих двух границ, число огибающих будет равно  
5 4, и устройство 100 может, снова, обратиться к следующему SBR фрейму. В качестве последнего шага, если никаких нарушений не обнаружено на границах 1, 2, 3, 5, 6, 7, детектор 130 может исследовать границу 4 и, если нарушение обнаружено на границе 4, число огибающих устанавливается на два.

10 Для общего случая ( $n$  временных частей, где  $n$  - четное число) эта процедура может также производиться следующим образом. Если, например, на нечетных границах какое-либо нарушение не обнаружено, и поэтому порог квантования 125 может быть ниже пороговой величины, что значит, что соседние огибающие (которые разделены этими границами) не включают серьезных различий относительно распределения  
15 спектральной энергии, то нет необходимости разделять SBR фрейм на  $n$  огибающих, а вместо этого может быть достаточно  $n/2$  огибающих. Если кроме того детектор 130 не обнаруживает нарушений на границах, которые являются дважды нечетным числом (например, на границах 2, 6, 10, ...), также нет необходимости помещать границу огибающей в эти положения и, следовательно, количество огибающих может далее  
20 быть уменьшено коэффициентом 2, то есть до  $n/4$ . Эта процедура продолжается шаг за шагом (следующим шагом будет граница, которая является четырежды нечетным числом, то есть 4, 12, ...). Если на всех этих границах какое-либо нарушение не обнаружено, будет достаточно единственной огибающей для целого SBR фрейма.

25 Если, однако, один из порогов квантования 125 на нечетных границах выше пороговой величины,  $n$  огибающие должны быть рассмотрены, так как только тогда граница огибающей будет помещена в соответствующее положение (так как предполагается, что все огибающие имеют ту же самую длину). В этом случае  $n$  огибающих будут вычисляться даже тогда, когда все другие пороги квантования 125  
30 ниже пороговой величины.

Детектор 130 может, однако, также рассмотреть все границы и рассмотреть все пороги квантования 125 для всех временных частей 110, чтобы вычислить число огибающих 104.

35 Так как увеличение числа огибающих 102 также подразумевает увеличенное количество данных, подлежащих передаче, порог принятия решения для соответствующей границы огибающей, который влечет за собой более высокое число огибающих 104, может быть увеличен. Это означает, что пороговая величина на границах 1, 3, 5 и 7 может, по выбору, быть выше, чем пороговая величина на  
40 границах 2 и 6, которая, в свою очередь, может быть выше, чем пороговая величина на границе 4. Более низкие или более высокие пороговые величины относятся к тому случаю, когда нарушение пороговой величины более или менее вероятно. Например, более высокая пороговая величина подразумевает, что отклонение в распределении спектральной энергии между двумя соседними временными частями может быть более  
45 допустимым, чем при более низкой пороговой величине, и, следовательно, для высокого порога необходимы более серьезные отклонения в распределении спектральной энергии, чтобы потребовались дальнейшие огибающие.

50 Выбранная пороговая величина может также зависеть от того, классифицирован ли сигнал как речевой сигнал или как обычный звуковой сигнал. Это, однако, не означает, что порог принятия решения будет всегда уменьшаться (или увеличиваться), если сигнал классифицируется как речь. В зависимости от применения, однако, может быть полезно, если для обычного звукового сигнала пороговая величина будет

высокая, в этом случае число огибающих обычно меньше, чем для речевого сигнала.

Фиг.5 иллюстрирует дальнейшие осуществления, в которых длина огибающих изменяется на протяжении SBR фрейма. На фиг.5а показан пример с тремя огибающими 104, первая огибающая 104а, вторая огибающая 104b и третья огибающая 104с. Первая огибающая 104а простирается от начального момента времени  $t_0$  до границы 2 в момент времени  $t_2$ , вторая огибающая 104b простирается от границы 2 в момент времени  $t_2$  до границы 5 в момент времени  $t_5$ , и третья огибающая 104с простирается от границы 5 в момент времени  $t_5$  до конечного момента времени  $t_n$ . Если опять все временные части имеют одну и ту же длину и если опять SBR фрейм разделен на восемь временных частей, то первая огибающая 104а покрывает первую и вторую временные части 111, 112, вторая огибающая 104b покрывает третью, четвертую и пятую временные части 113-115 и третья огибающая 104с покрывают шестую, седьмую и восьмую временные части. Поэтому первая огибающая 104а меньше, чем вторая и третья огибающие 104b и 104с.

Фиг.5b показывает другое осуществление только с двумя огибающими; первая огибающая 104а простирается от начального момента времени  $t_0$  до первого момента времени  $t_1$ , и вторая огибающая 104b простирается от первого момента времени  $t_1$  до конечного момента времени  $t_n$ . Поэтому вторая огибающая 104b простирается по 7-ми временным частям, тогда как первая огибающая 104а простирается только по единственной временной части (первая временная часть 111).

Фиг.5с снова показывает осуществление с тремя огибающими 104, где первая огибающая 104а простирается от начального момента времени  $t_0$  до второго момента времени  $t_2$ , вторая огибающая 104b простирается от второго момента времени  $t_2$  до четвертого момента времени  $t_4$ , и третья огибающая 104с простирается от четвертого момента времени  $t_4$  до конечного момента времени  $t_n$ .

Эти осуществления могут, например, использоваться в случае, когда границы огибающих 104 помещены только между соседними временными частями, в которых нарушение пороговой величины обнаруживается или в начальном моменте времени  $t_0$ , или в конечном моменте времени  $t_n$ . Это означает, что на фиг.5а нарушение обнаруживается в моменте времени  $t_2$ , и нарушение обнаруживается в моменте времени  $t_5$ , тогда как никакие нарушения не обнаруживаются в остающиеся моменты времени  $t_1$ ,  $t_3$ ,  $t_4$ ,  $t_6$  и  $t_7$ . Точно так же на фиг.5b нарушение обнаруживается только в момент времени  $t_1$ , в результате чего появляется граница для первой огибающей 104а и для второй огибающей 104b, а на фиг.5с нарушение обнаруживается только во втором моменте времени  $t_2$  и в четвертом моменте времени  $t_4$ .

Чтобы декодер мог использовать данные огибающей и копировать соответственно верхнюю спектральную полосу, декодеру требуется положение огибающей 104 и соответствующих границ огибающей. Как показано ранее, в осуществлениях, которые опираются на указанный стандарт, где все огибающие 104 имеют ту же самую длину, и, следовательно, этого достаточно, чтобы передать число огибающих так, чтобы декодер мог решить, где должна быть граница огибающей. В этих осуществлениях, как показано на фиг.5, однако, декодеру требуется информация о том, в какой момент времени помещена граница огибающей, и таким образом дополнительная информация может быть введена в поток данных так, чтобы, используя дополнительную информацию, декодер мог сохранить моменты времени, куда помещена граница и начала и концы огибающих. Эта дополнительная информация включает момент времени  $t_2$  и  $t_5$  (случай на фиг.5а), момент  $t_1$  (случай на фиг.5b) и момент времени  $t_2$  и  $t_4$  (случай на фиг.5с).

Фиг.6А и 6В показывают осуществление для вычислителя порога квантования 120 посредством использования распределения спектральной энергии в звуковом сигнале 105.

5 Фиг.6А показывает первое множество выборочных значений 610 для звукового сигнала в данной временной части, например первая временная часть 111, и сравнивает этот выбранный звуковой сигнал со вторым множеством образцов звукового сигнала 620 во второй временной части 112. Звуковой сигнал был преобразован в частотную область так, чтобы множества выборочных значений 610, 10 620 или их уровни  $P$  были показаны как функция частоты  $f$ . Нижний и верхний частотные диапазоны отделяются частотой разделения  $f_0$ , подразумевая, что для частот выше  $f_0$  выборочные значения передаваться не будут. Вместо этого декодер должен копировать эти выборочные значения посредством использования данных SBR. С другой стороны, образцы, ниже частоты разделения  $f_0$ , кодируются, 15 например, кодирующим устройством AAC и передаются декодеру.

Декодер может использовать эти выборочные значения из низкочастотного диапазона, чтобы копировать высокочастотные компоненты. Поэтому, чтобы найти меру отклонения первого множества образцов 610 в первой временной части 111 и 20 второго множества образцов 620 во второй временной части 112, может быть недостаточно рассмотреть только выборочные значения в высокочастотном диапазоне (для  $f > f_0$ ), но также и принять во внимание частотные компоненты в низкочастотном диапазоне. В общем можно ожидать хорошее качество репликации, если имеется корреляция между частотными компонентами в высокочастотном 25 диапазоне с частотными компонентами в низкочастотном диапазоне. На первом этапе этого может быть достаточно, чтобы рассмотреть только выборочные значения в высокочастотном диапазоне (выше частоты разделения  $f_0$ ) и чтобы вычислить корреляцию между первым множеством выборочных значений 610 со вторым 30 множеством выборочных значений 620.

Корреляция может быть вычислена при использовании стандартных статистических способов и может включать, например, вычисление так называемой функции взаимной корреляции или других статистических мер подобия двух сигналов. Существует также коэффициент корреляции смешанного момента Пирсона, который 35 может использоваться, чтобы оценивать корреляцию двух сигналов. Коэффициенты Пирсона также известны как выборочные коэффициенты корреляции. В общем корреляция указывает силу и направление линейного отношения между двумя случайными переменными - в этом случае, - двумя выборочными распределениями 610 40 и 620. Поэтому корреляция относится к отклонению двух случайных переменных от независимости. В широком смысле имеется несколько коэффициентов, измеряющих степень корреляции, приспособленных к характеру данных так, чтобы различные коэффициенты использовались для различных ситуаций.

Фиг.6В показывает третье множество выборочных значений 630 и четвертое 45 множество выборочных значений 640, которое может, например, быть связано с выборочными значениями в третьей временной части 113 и в четвертой временной части 114. Снова, чтобы сравнить два множества образцов (или сигналов), рассматривают две соседние временные части. В отличие от случая, показанного на 50 фиг.6А, на фиг.6В введена пороговая величина  $T$  так, чтобы рассматривались только выборочные значения, чей уровень  $P$  выше (или более часто нарушает) пороговой величины  $T$  (для которой поддерживается  $P > T$ ).

В этом осуществлении отклонение в распределениях спектральной энергии может

быть измерено просто посредством подсчета числа выборочных значений с нарушением этой пороговой величины  $T$ , и в результате может быть установлен порог квантования 125. Этот простой способ приведет к корреляции между обоими сигналами без проведения детального статистического анализа различных множеств выборочных значений в различных временных частях 110. Альтернативно, статистический анализ, например, как упомянуто выше, может применяться к образцам, которые нарушают только пороговую величину  $T$ .

Фиг.7А-7С показывают дальнейшее осуществление, где кодирующее устройство 300 включает переключающийся блок принятия решения 370 и блок стереокодирования 380. Кроме того, кодирующее устройство 300 также включает инструменты расширения полосы пропускания такие, как, например, вычислитель данных огибающей 210 и SBR-связанные модули 310. Переключающийся блок принятия решения 370 обеспечивает сигнал переключения принятия решений 371, который переключается между звуковым кодирующим устройством 372 и речевым кодирующим устройством 373. Каждый из этих кодов может кодировать звуковой сигнал в основном частотном диапазоне, используя различные числа выборочных значений (например, 1024 для более высокого разрешения или 256 для более низкого разрешения). Сигнал переключения принятия решений 371 также поставляется инструменту расширения полосы пропускания (BWE) 210, 310. Инструмент BWE 210, 310 будет тогда использовать переключение принятия решений 371 чтобы, например, приспособить пороговые величины для определения числа 102 огибающих спектра 104 и включить / выключить дополнительный детектор кратковременных помех. Звуковой сигнал 105 вводится в переключающийся блок принятия решений 370 и вводится в стереокодирование 380 так, чтобы стереокодирование 380 могло произвести выборочные значения, которые вводятся в блок расширения полосы пропускания 210, 310. В зависимости от решения 371, генерированного переключающимся блоком принятия решений 370, инструмент расширения полосы пропускания 210, 310 будет генерировать данные репликации спектрального диапазона, которые, в свою очередь, направляются или звуковому кодирующему устройству 372, или речевому кодирующему устройству 373.

Сигнал переключения принятия решений 371 зависит от сигнала, и может быть получен переключающимся блоком принятия решений 370 посредством анализа звукового сигнала, например, при использовании детектора кратковременных помех или других детекторов, которые могут включать или не включать переменную пороговую величину. Альтернативно, сигнал переключения принятия решений 371 может также приспособливаться вручную или получаться из потока данных (включенного в звуковой сигнал).

Выход звукового кодирующего устройства 372 и речевого кодирующего устройства 373 может снова быть введен в форматер битового потока 350 (см. фиг.3А).

Фиг.7В показывает пример сигнала переключения принятия решений 371, который обнаруживает звуковой сигнал для промежутка времени ниже первого момента времени  $t_a$  и выше второго момента времени  $t_b$ . Между первым моментом времени  $t_a$  и вторым моментом времени  $t_b$  переключающийся блок принятия решений 370 обнаруживает речевой сигнал, подразумевающий различные дискретные значения для сигнала переключения принятия решения 371.

В результате, как показано на фиг.7С, в течение времени обнаруживается звуковой сигнал, что означает, что для моментов времени до  $t_a$  временное разрешение кодирования будет низким, тогда как во временном промежутке, где обнаруживается

речевой сигнал (между первым моментом времени  $t_a$  и вторым моментом  $t_b$ ), временное разрешение увеличивается. Увеличение временного разрешения подразумевает более короткое окно анализа во временном интервале. Увеличенное временное разрешение подразумевает также вышеупомянутое увеличенное число огибающих спектра (см. описание к фиг.4).

Для речевых сигналов, требующих точного временного представления высоких частот, порог принятия решения (например, используемый на фиг.4) для передачи более высокого числа множеств значений параметра управляется переключающимся блоком принятия решения 370. Для речевых и речеподобных сигналов, закодированных кодирующей частью речи или временного интервала 373 переключающегося основного кодирующего устройства, порог принятия решения для использования больших множеств значений параметра может, например, быть снижен и, поэтому, временное разрешение увеличивается. Это, однако, не всегда так, как упомянуто выше. Адаптация времениподобного разрешения к сигналу не зависит от базовой структуры кодирующего устройства (который не использовался на фиг.4). Это означает, что описанный способ также применим для системы, в которой модуль SBR включает только одиночное основное кодирующее устройство.

Хотя некоторые аспекты были описаны в контексте устройства, ясно, что эти аспекты также представляют описание соответствующего способа, где блок или устройство соответствуют этапу способа или характерной особенности этапа способа. Аналогично, аспекты, описанные в контексте этапа способа, также представляют описание соответствующего блока или пункта или характерной особенности соответствующего устройства.

Изобретательный закодированный звуковой сигнал может храниться на цифровом носителе данных или может быть передан по каналам передачи, таким как беспроводные каналы передачи или проводные каналы передачи, такие как Интернет.

В зависимости от определенных требований выполнения осуществления изобретения могут быть исполнены в аппаратных средствах или в программном обеспечении. Выполнение может быть осуществлено при использовании цифрового носителя данных, например дискета, DVD, компакт-диск, ROM (постоянное запоминающее устройство, ПЗУ), PROM (программируемое постоянное запоминающее устройство, ППЗУ), EPROM (стираемое программируемое постоянное запоминающее устройство СППЗУ), EEPROM (электрически стираемое программируемое постоянное запоминающее устройство, ЭСППЗУ), или флэш-память, имеющих сохраненные на нем электронно-считываемые управляющие сигналы, которые взаимодействуют (или могут взаимодействовать) с программируемой вычислительной системой таким образом, что реализуется соответствующий способ.

Некоторые осуществления согласно изобретению включают носитель информации, имеющий электронно-считываемые управляющие сигналы, которые способны взаимодействовать с программируемой вычислительной системой таким образом, что реализуется один из описанных здесь способов.

В общем осуществления данного изобретения могут быть выполнены как компьютерный программный продукт с управляющей программой; управляющая программа служит для выполнения одного из способов, когда компьютерный программный продукт запущен на компьютере. Управляющая программа может, например, сохраняться на машиночитаемом носителе.

Другие осуществления включают компьютерную программу для реализации

одного из описанных здесь способов, сохраненную на машиночитаемом носителе.

Другими словами, осуществлением предлагаемого способа, поэтому, является компьютерная программа, имеющая управляющую программу для реализации одного из описанных здесь способов, когда компьютерная программа запущена на компьютере.

Дальнейшим осуществлением предлагаемых способов, поэтому, является носитель информации (или цифровой носитель информации, или считываемая компьютером информация), включающий записанную на нем компьютерную программу для реализации одного из описанных здесь способов.

Дальнейшим осуществлением предлагаемого способа, поэтому, является поток данных или последовательность сигналов, представляющих компьютерную программу для реализации одного из описанных здесь способов. Поток данных или последовательность сигналов могут, например, формироваться, чтобы быть переданными через канал передачи данных, например через Интернет.

Дальнейшее осуществление включает средство обработки, например компьютер, или программируемое логическое устройство, формируемое для или приспособленное для выполнения одного из описанных здесь способов.

Дальнейшее осуществление включает компьютер с установленной на нем компьютерной программой для реализации одного из описанных здесь способов.

В некоторых осуществлениях программируемое логическое устройство (например, промышленная программируемая логическая матрица) может использоваться для выполнения некоторых или всех функциональных возможностей описанных здесь способов. В некоторых осуществлениях промышленная программируемая логическая матрица может взаимодействовать с микропроцессором, чтобы выполнить один из описанных здесь способов. Обычно способы предпочтительно выполняются любым аппаратным оборудованием.

Вышеописанные осуществления являются только иллюстрацией принципов данного изобретения. Имеется в виду, что модификации и изменения расположения и деталей, описанных здесь, будут понятны специалистам, квалифицированным в этой области. Поэтому целью является то, чтобы ограничиваться только объемом формулы изобретения, а не специфическими деталями, представленными посредством приведенных здесь описаний и объяснений осуществлений.

### Формула изобретения

1. Устройство (100) для вычисления числа (102) огибающих спектра (104), которые должны быть получены посредством кодирующего устройства для репликации спектральной полосы (SBR), где кодирующее устройство SBR приспособлено, чтобы кодировать звуковой сигнал (105), используя множество выборочных значений в пределах предварительно определенного числа последующих временных частей (110) в SBR фрейме, простирающемся от начального момента времени ( $t_0$ ) до конечного момента времени ( $t_n$ ); предварительно определенное число последующих временных частей (110) расположено во временной последовательности, предоставленной звуковым сигналом (105); устройство (100) включает: вычислитель порога квантования (120) для определения порога квантования (125); порог квантования (125) измеряет отклонения в распределениях спектральной энергии пары соседних временных частей; детектор (130) для обнаружения нарушения (135) пороговой величины при помощи порога квантования (125); процессор (140) для определения первой границы огибающей (145) между парой соседних временных частей, когда

обнаружено нарушение (135) пороговой величины; процессор (150) для определения второй границы огибающей (155) между отличной парой соседних временных частей, или в начальный момент времени ( $t_0$ ), или в конечный момент времени ( $t_n$ ) для огибающей, имеющей первую границу огибающей (145), основанную на  
5 нарушении (135) пороговой величины для другой пары, или основанную на временном положении пары или отличной пары в SBR фрейме; и числовой процессор (160) для установления числа (102) огибающих спектра (104), имеющих первую границу огибающей (145) и вторую границу огибающей (155), где  
10 предварительно определенное число временных частей (110) равно  $n$  с  $n-1$  границами между соседними временными частями (110), которые пронумерованы и организованы относительно времени так, чтобы границы включали четные и нечетные границы, и где числовой процессор (160) приспособлен, чтобы установить  $n$ , как число (102) огибающих спектра (104), если детектор (130) обнаруживает  
15 нарушение (135) на нечетной границе, где детектор (150) приспособлен, чтобы определить вторую границу (155) таким образом, чтобы огибающие спектра (104) включали ту же самую продолжительность и число (102) огибающих спектра (104), равное степени двух или, где устройство (100) далее включает переключающийся блок  
20 принятия решения (370), формируемый, чтобы предоставить сигнал переключения принятия решения (371); сигнал переключения принятия решения (371) подает речеподобный звуковой сигнал и обычный звукоподобный звуковой сигнал, где детектор (130) приспособлен, чтобы понижать пороговую величину для речеподобных звуковых сигналов.

2. Устройство (100) по п.1, в котором длительность временной части  
25 предварительно определенного числа последующих временных частей (110) равна минимальной продолжительности, для которой определена единственная огибающая, и в котором вычислитель порога квантования (120) приспособлен для вычисления порога квантования (125) для двух соседних временных частей, имеющих  
30 минимальную продолжительность.

3. Устройство (100) по п.1, где процессор (140) приспособлен, чтобы установить первую границу (145) в первом обнаруженном нарушении (135), и где процессор (150) приспособлен, чтобы установить вторую границу огибающей (155) после сравнения,  
35 по крайней мере, одного другого порога квантования (125) с пороговой величиной.

4. Устройство (100) по п.3, далее включает информационный процессор для предоставления дополнительной информации; дополнительная информация включает первую границу огибающей (145) и вторую границу огибающей (155) в пределах  
40 временной последовательности звукового сигнала (105).

5. Устройство (100) по п.1, где детектор (130) приспособлен, чтобы исследовать во временной последовательности каждую из границ между соседними временными частями (110).

6. Устройство (100) по п.6, где детектор (130) приспособлен, чтобы обнаружить  
45 первое нарушение (135) на нечетных границах.

7. Устройство (100) по п.8, где предварительно определенное число равно 8, и где числовой процессор (160) приспособлен, чтобы установить число (102) огибающих спектра (104) на 1, 2, 4 или 8 таким образом, чтобы каждая из огибающих спектра (104) включала ту же самую продолжительность.  
50

8. Устройство (100) по п.1 или 7, где детектор (130) приспособлен, чтобы использовать пороговую величину, которая зависит от временного положения нарушения (135), таким образом, что во временном положении, производящем

большее число огибающих спектра (104), используется более высокая пороговая величина, чем для временного положения, производящего более низкое число огибающих спектра (104).

5 9. Устройство (100) по п.1 далее включает детектор кратковременных помех с переходной пороговой величиной; переходная пороговая величина больше, чем пороговая величина, и/или далее включает вычислитель данных огибающей (210); вычислитель данных огибающей (210) приспособлен, чтобы вычислять данные огибающей спектра для огибающей спектра (104), распространяющейся от первой  
10 границы огибающей (145) до второй границы огибающей (155).

10. Кодировующее устройство (300) для кодирования звукового сигнала (105) включает:

основное кодировующее устройство (340) для кодирования звукового сигнала (105) в пределах основного частотного диапазона; устройство (100) для вычисления  
15 числа (102) огибающих спектра (104) по одному из пп.1-9; и вычислитель данных огибающей (210) для вычисления данных огибающей, основанных на звуковом сигнале (105) и числе (102).

11. Способ вычисления числа (102) огибающих спектра (104), которые должны быть  
20 получены кодировующим устройством для репликации спектральной полосы (SBR), где кодировующее устройство SBR приспособлено, чтобы кодировать звуковой сигнал (105), используя множество выборочных значений в пределах предварительно определенного числа последующих временных частей (110) в SBR фрейме, простирающемся от начального момента времени ( $t_0$ ) до конечного момента  
25 времени ( $t_n$ ); предварительно определенное число последующих временных частей (110) организуется во временной последовательности, предоставленной звуковым сигналом (105); способ включает: определение порога квантования (125); порог квантования (125) измеряет отклонение в распределениях спектральной энергии пары соседних временных частей; обнаружение нарушения (135) пороговой величины при помощи порога квантования (125); определение первой границы огибающей (145) между парой соседних временных частей, когда обнаружено нарушение (135) пороговой величины; определение второй границы огибающей (155) между отличной парой соседних временных частей, или в начальный момент времени ( $t_0$ ) или в  
35 конечный момент времени ( $t_n$ ) для огибающей, имеющей первую границу огибающей (145), основанной на нарушении (135) пороговой величины для другой пары, или основанной на временном положении пары или отличной пары в SBR фрейме; и установление числа (102) огибающих спектра (104), имеющих первую  
40 границу огибающей (145) и вторую границу огибающей (155), где предварительно определенное число временных частей (110) равно  $n$  с  $n-1$  границами между соседними временными частями (110), которые пронумерованы и организованы относительно времени так, чтобы границы включали четные и нечетные границы, и где числовой процессор (160) приспособлен, чтобы установить  $n$ , как число (102) огибающих  
45 спектра (104), если детектор (130) обнаруживает нарушение (135) на нечетной границе или, где детектор (150) приспособлен, чтобы определить вторую границу (155) таким образом, чтобы огибающие спектра (104) включали ту же самую продолжительность и число (102) огибающих спектра (104), равное степени двух или, где устройство (100) далее включает переключающийся блок принятия решения (370), формируемый, чтобы  
50 предоставить сигнал переключения принятия решения (371); сигнал переключения принятия решения (371) подает речеподобный звуковой сигнал и обычный звукоподобный звуковой сигнал, где детектор (130) приспособлен, чтобы понижать

пороговую величину для речеподобных звуковых сигналов.

12. Машиночитаемый носитель информации с сохраненной на нем компьютерной программой, имеющей код для реализации способа по п.11, когда программа запущена на процессоре.

5

10

15

20

25

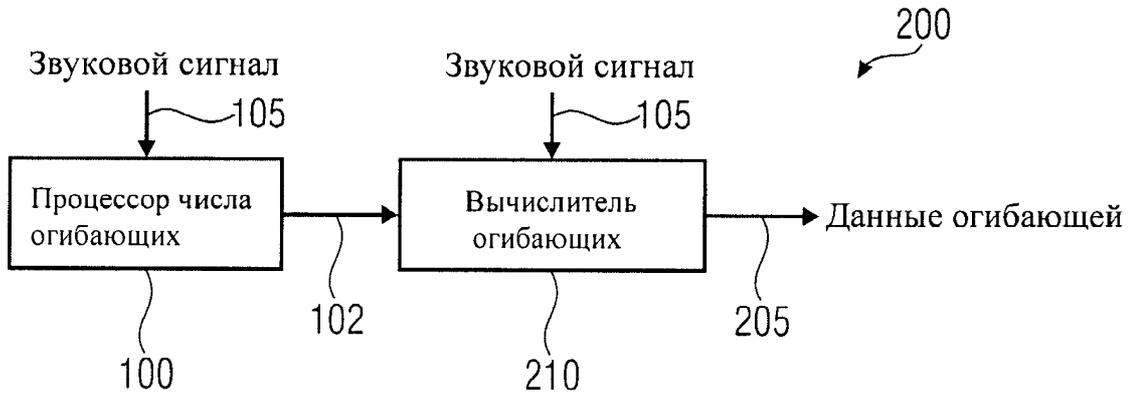
30

35

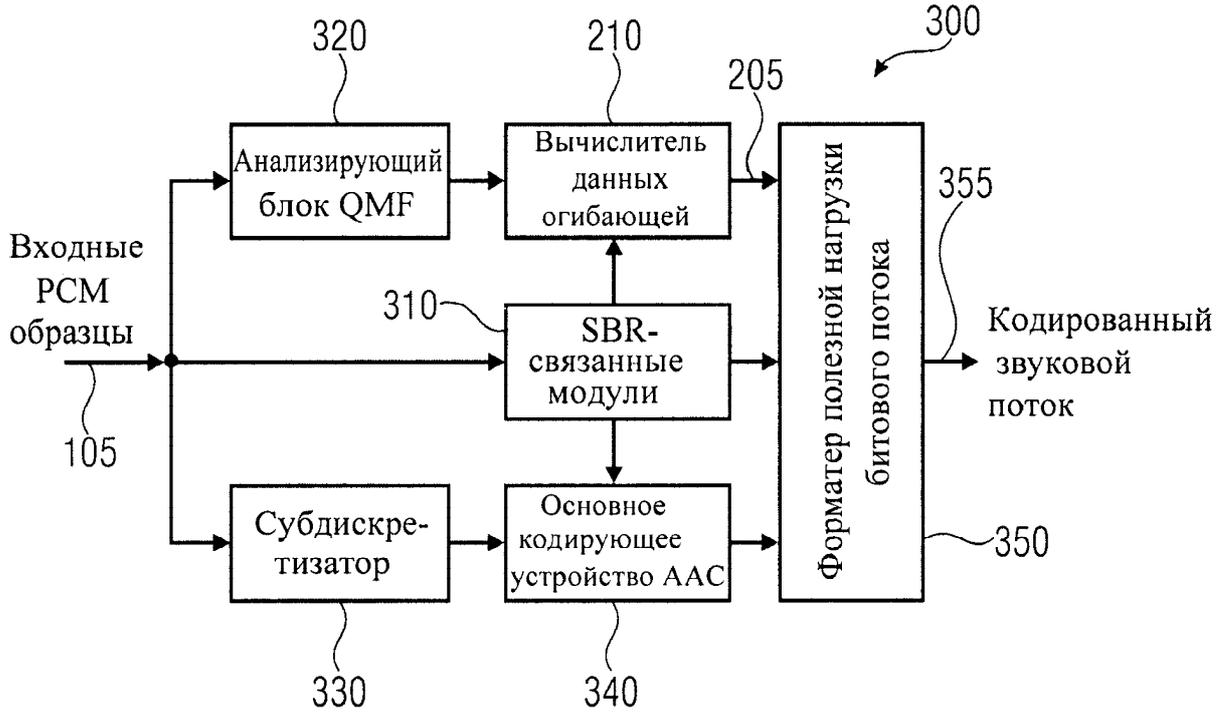
40

45

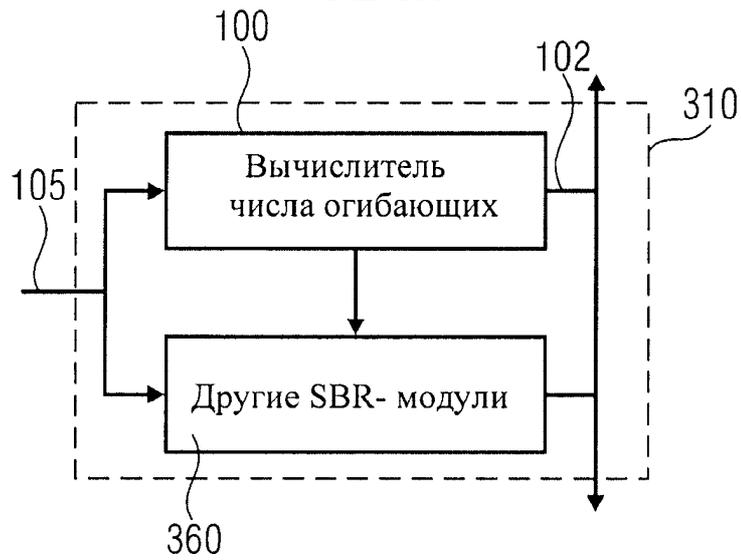
50



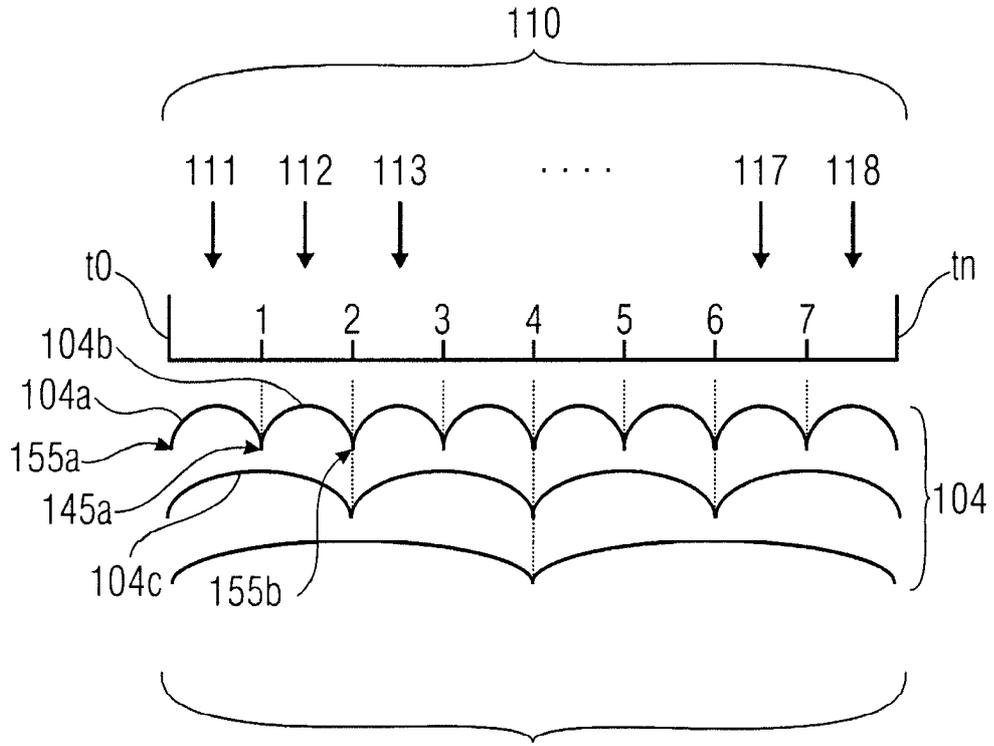
Фиг. 2



Фиг. 3А

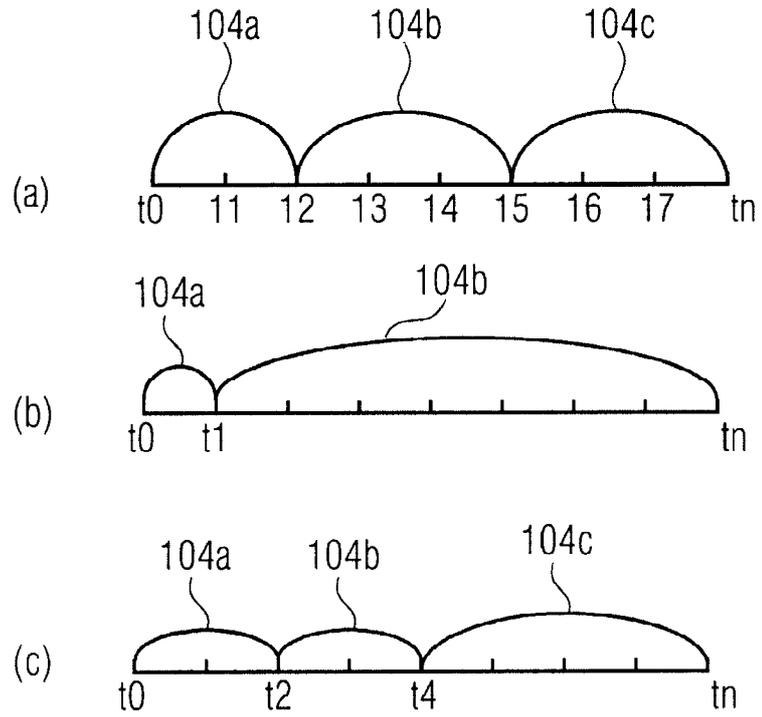


Фиг. 3В

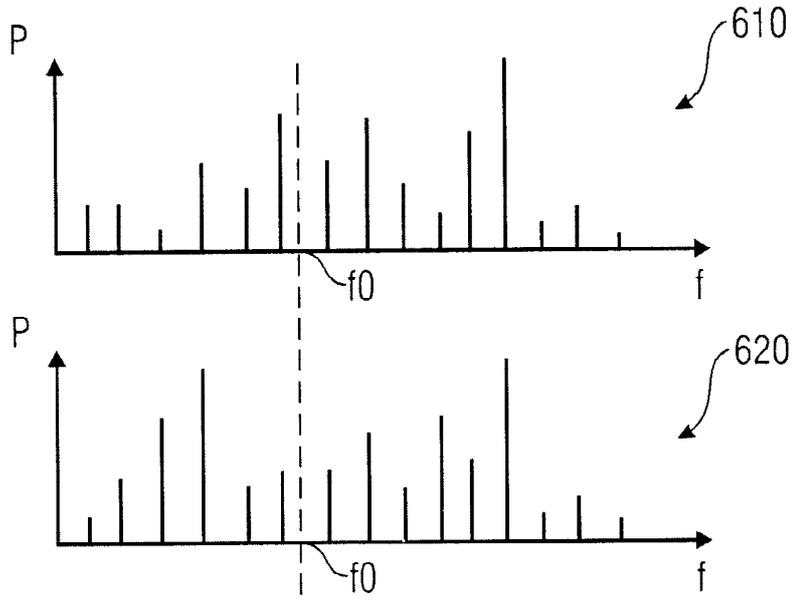


SBR- структура

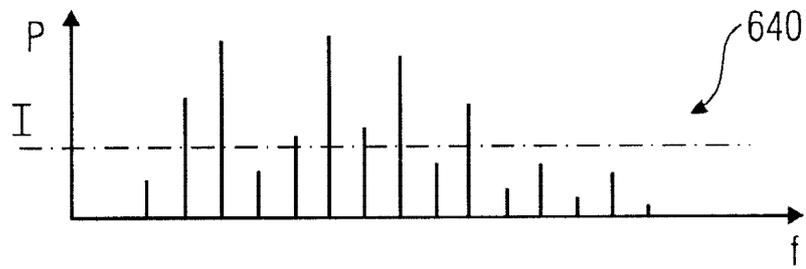
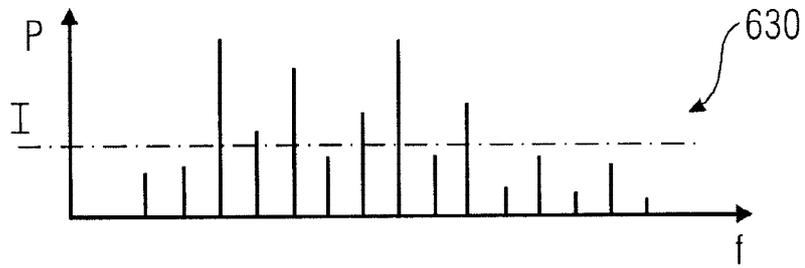
Фиг. 4



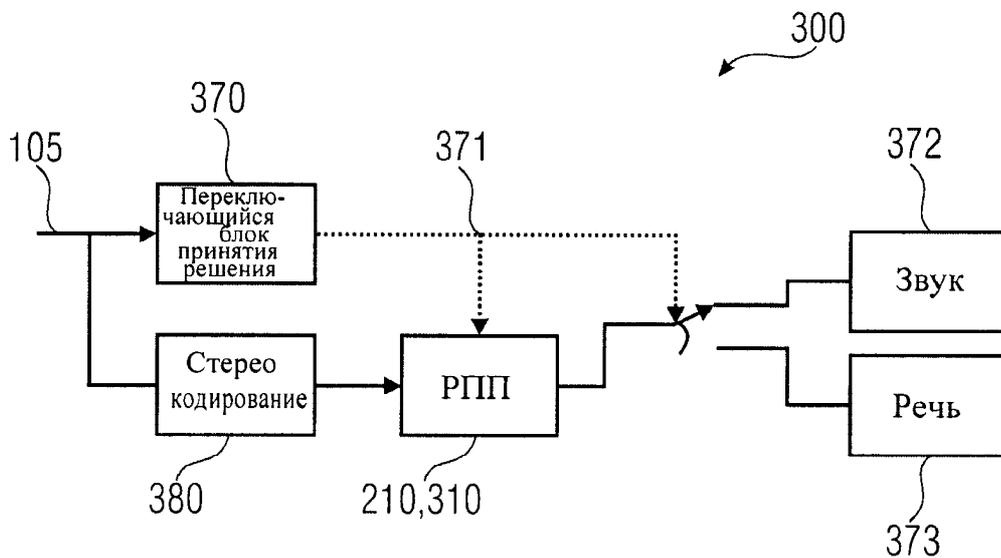
Фиг. 5



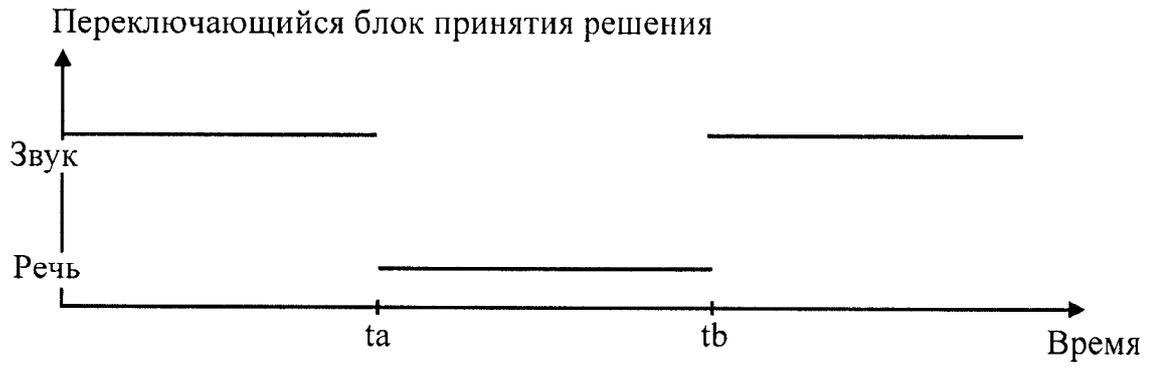
Фиг. 6А



Фиг. 6В

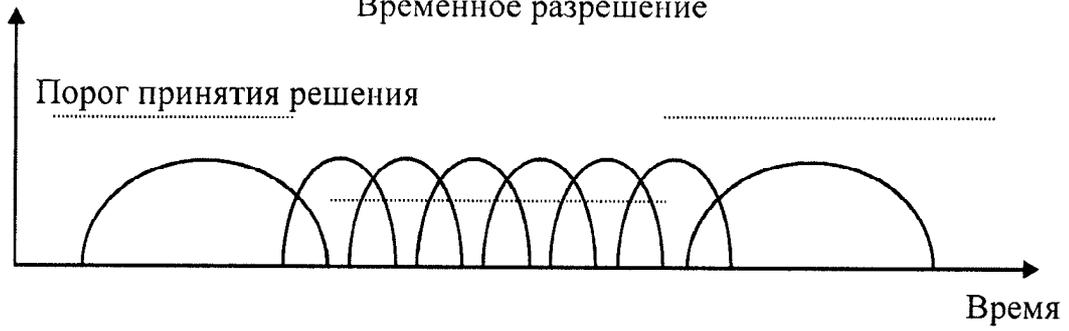


Фиг. 7А



Фиг. 7В

Временное разрешение



Фиг. 7С