



(19) **RU** <sup>(11)</sup> **2 236 046** <sup>(13)</sup> **C2**  
 (51) МПК<sup>7</sup> **G 10 L 19/00//G 10 L 101:023**

РОССИЙСКОЕ АГЕНТСТВО  
 ПО ПАТЕНТАМ И ТОВАРНЫМ ЗНАКАМ

**(12) ОПИСАНИЕ ИЗОБРЕТЕНИЯ К ПАТЕНТУ РОССИЙСКОЙ ФЕДЕРАЦИИ**

(21), (22) Заявка: 2002111665/09, 29.09.2000  
 (24) Дата начала действия патента:  
 29.09.2000ii.1, 17-19  
 (30) Приоритет: 01.10.1999 SE 9903552-9ii.2-16  
 26.01.2000 SE PCT/SE00/00158  
 (43) Дата публикации заявки: 10.11.2003  
 (46) Дата публикации: 10.09.2004  
 (56) Ссылки: US 5852806 A, 22.12.1998. US 5623577  
 A, 22.04.1997. RU 96121146 C1, 27.01.1999.  
 US 5581653 A, 03.12.1996. US 5737718 A,  
 07.04.1998.  
 (85) Дата перевода заявки PCT на национальную  
 фазу: 06.05.2002  
 (86) Заявка PCT:  
 SE 00/01887 (29.09.2000)  
 (87) Публикация PCT:  
 WO 01/26095 (12.04.2001)  
 (98) Адрес для переписки:  
 129010, Москва, ул. Б. Спасская, 25, стр.3,  
 ООО "Юридическая фирма Городисский и  
 Партнеры", пат.пов. Ю.Д.Кузнецову, рег.№ 595

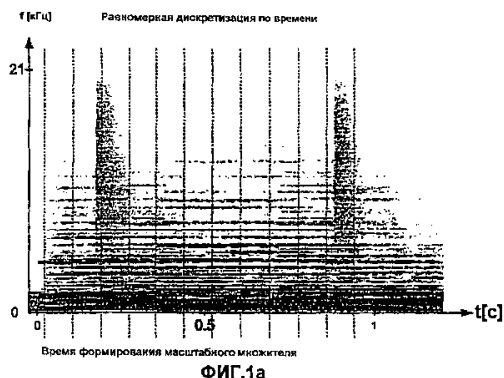
(72) Изобретатель: ЛИЛЬЕРЮД Ларс Густаф (SE),  
 КЬЕРЛИНГ Кристофер (SE), ЭКСТРАНД Пер  
 (SE), ХЕНН Фредрик (SE)  
 (73) Патентообладатель:  
 КОУДИНГ ТЕКНОЛОДЖИЗ СВИДН АБ (SE)  
 (74) Патентный поверенный:  
 Кузнецов Юрий Дмитриевич

RU 2 236 046 C 2

RU 2 236 046 C 2

**(54) ЭФФЕКТИВНОЕ КОДИРОВАНИЕ ОГИБАЮЩЕЙ СПЕКТРА С ИСПОЛЬЗОВАНИЕМ ПЕРЕМЕННОГО РАЗРЕШЕНИЯ ПО ВРЕМЕНИ И ПО ЧАСТОТЕ И ПЕРЕКЛЮЧЕНИЯ ВРЕМЯ/ЧАСТОТА**

(57)  
 Изобретения относятся к кодированию аудиосигналов и могут быть использованы для передачи сигналов. Техническим результатом является исключение избыточности при передаче сигналов. Изобретения основаны на передаче сигналами частотно-временного отображения представления огибающей спектра на основе адаптивного частотно-временного направленного кодирования. 4 с. и 15 з.п. ф-лы, 12 ил.





(19) **RU** <sup>(11)</sup> **2 236 046** <sup>(13)</sup> **C2**  
 (51) Int. Cl.<sup>7</sup> **G 10 L 19/00//G 10 L 101:023**

RUSSIAN AGENCY  
 FOR PATENTS AND TRADEMARKS

(12) **ABSTRACT OF INVENTION**

(21), (22) Application: 2002111665/09, 29.09.2000

(24) Effective date for property rights:  
 29.09.2000 с 1.1.17-19

(30) Priority: 01.10.1999 SE 9903552-9 с 1.2-16  
 26.01.2000 SE PCT/SE00/00158

(43) Application published: 10.11.2003

(46) Date of publication: 10.09.2004

(85) Commencement of national phase: 06.05.2002

(86) PCT application:  
 SE 00/01887 (29.09.2000)

(87) PCT publication:  
 WO 01/26095 (12.04.2001)

(98) Mail address:  
 129010, Moskva, ul. B. Spasskaja, 25, str.3,  
 ООО "Juridicheskaja firma Gorodisskij i  
 Partnery", pat.pov. Ju.D.Kuznetsovu, reg.№ 595

(72) Inventor: LIL'ERJuD Lars Gustaf (SE),  
 K'ERLING Kristofer (SE), EhKSTRAND Per  
 (SE), KhENN Fredrik (SE)

(73) Proprietor:  
 KOUDING TEKNOLODZhIZ SVIDN AB (SE)

(74) Representative:  
 Kuznetsov Jurij Dmitrievich

(54) EFFECTIVE ENCODING OF SPECTRUM ENVELOPE WITH USE OF VARIABLE RESOLUTION IN TIME AND FREQUENCY AND SWITCHING TIME/FREQUENCY

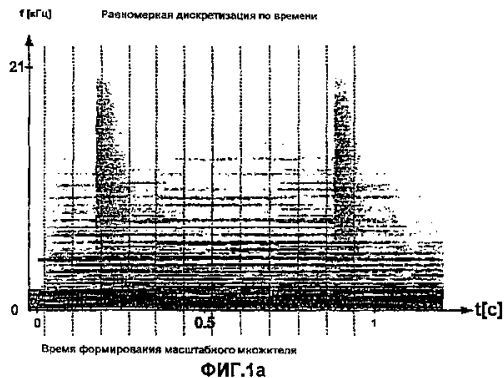
(57) Abstract:

FIELD: audio signals encoding technology.

SUBSTANCE: method is based on transfer of frequency-time image of spectrum envelope on basis of adaptive time-frequency directed encoding.

EFFECT: prevented superfluity during signals transfer.

12 dwg



RU 2 236 046 C2

RU 2 236 046 C2

Область техники, к которой относится изобретение

Настоящее изобретение относится к новому способу и устройству эффективного кодирования огибающих спектра в системах кодирования аудиосигналов. Способ может быть использован как для кодирования естественных аудиосигналов, так и для кодирования речи и, главным образом, предназначен для кодеров, использующих дублирование полосы спектра (WO 98/57436) или другие способы высокочастотного восстановления.

Уровень техники

Методы кодирования источников аудиосигналов могут быть разделены на два класса: кодирование естественных аудиосигналов и кодирование речи. Кодирование естественных аудиосигналов обычно используется для музыкальных или произвольных сигналов со средними скоростями передачи битов и, как правило, занимает широкую полосу звуковых частот. Речевые кодеры, в основном, ограничиваются воспроизведением речи, но, с другой стороны, могут быть использованы при очень низких скоростях передачи битов, хотя и с узкой полосой звуковых частот. В обоих случаях сигнал обычно разделяется на две основные составляющие сигнала - "огибающую спектра" и соответствующий сигнал "остатка". В нижеследующем описании термин "огибающая спектра" относится к грубому спектральному распределению сигнала в общем смысле, например коэффициенты фильтра в кодере, основанном на линейном предсказании, или множество частотно-временных средних значений выборки поддиапазонов в многополосном кодере. Термин "остаток" относится к точному спектральному распределению в общем смысле, например сигнал ошибки кодирования с линейным предсказанием или выборки поддиапазонов, нормированные с использованием вышеуказанных частотно-временных средних значений. "Данные огибающей" относятся к квантованной и кодированной огибающей спектра, а "данные остатка" - к квантованному и кодированному остатку. При средних и высоких скоростях передачи битов данные остатка составляют основную часть битового потока. При очень низких скоростях передачи битов данные огибающей составляют большую часть битового потока. Следовательно, действительно важно компактно представить огибающую спектра при использовании более низких скоростей передачи битов.

Известные кодеры аудиосигналов и большинство речевых кодеров используют относительно короткие временные сегменты с постоянной длиной при формировании данных огибающей, чтобы получить хорошее разрешение во времени. Это предотвращает, однако, оптимальное использование маскирования в частотной области, известного из психоакустики. Для того чтобы повысить выигрыш от кодирования посредством использования узких полос фильтра с крутыми спадами и тем не менее достичь хорошего разрешения по времени во время переходных процессов, в современных кодерах аудиосигналов используется адаптивное переключение окна, т.е. в них

переключается длина временных сегментов в зависимости от статистических характеристик сигнала. Ясно, что минимальное использование коротких сегментов является предпосылкой для максимального выигрыша от кодирования. К сожалению, длинные окна переходов необходимы для изменения длины сегментов, но они ограничивают гибкость переключения.

Огибающая спектра представляет собой функцию двух переменных: времени и частоты. Кодирование может быть выполнено с использованием избыточности в обоих направлениях частотно-временной плоскости. Обычно кодирование огибающей спектра выполняется в частотном направлении с использованием дельта-кодирования (дифференциальной импульсно-кодовой модуляции (ДИКМ)) или векторного квантования (ВК).

Сущность изобретения

Настоящее изобретение предлагает новый способ и устройство кодирования огибающей спектра. Схема кодирования предназначена для выполнения особых требований систем, в которых сигнал остатка в некоторых областях частот исключается из передаваемых данных. Примерами являются системы, использующие высокочастотное восстановление (ВЧВ), в частности дублирование полосы спектра (ДПС) или параметрические кодеры. При одном исполнении неравномерная дискретизация по времени и по частоте огибающей спектра достигается посредством адаптивного группирования выборки поддиапазонов от банка фильтров фиксированного размера в полосы частот и временные сегменты, каждый из которых формирует одну выборку огибающей. Это позволяет производить мгновенный выбор произвольного разрешения по времени и по частоте в пределах банка фильтров. По умолчанию система установлена на длинные временные сегменты и высокое разрешение по частоте. Вблизи переходных процессов используются более короткие временные сегменты, посредством чего могут быть использованы большие шаги частоты для сохранения в заданных пределах размера данных. Для того чтобы максимизировать преимущества неравномерной дискретизации по времени, используется переменная длина фреймов (кадров) или гранул битового потока. Способ с переменным разрешением по времени и по частоте также применим к кодированию огибающей на основании предсказания. Вместо группирования выборки поддиапазонов формируются коэффициенты предсказания для временных сегментов переменной длины согласно системе.

В изобретении описываются две схемы для передачи сигналов используемого разрешения по времени и по частоте. Первая схема позволяет выполнять произвольный выбор посредством явной передачи сигналами границ временных сегментов и разрешения по частоте. Для того чтобы уменьшить количество служебных сигналов, используется четыре класса гранул, при этом предлагаются различные оптимальные соотношения затраты/гибкость. Во второй схеме используется свойство типичной программы, заключающееся в том, что переходные процессы разделяются, по меньшей мере, временным промежутком

$T_{nmin}$ , чтобы дополнительно уменьшить количество битов управления. Таким образом, детектор переходных процессов в кодере, работающий во временном интервале  $T_{det} \leq T_{nmin}$ , равном номинальной длине гранулы, определяет положение начала возможного переходного процесса. Положение внутри интервала кодируется и посылается на декодер. Кодер и декодер совместно используют правила, которые точно определяют частотно-временное распределение выборок огибающей спектра при наличии некоторой комбинации последовательных сигналов управления, обеспечивая однозначное декодирование данных огибающей.

Настоящее изобретение предлагает новый и эффективный способ избыточного кодирования с масштабными множителями. Импульс Дирака во временной области преобразуется в постоянную в частотной области, и функция Дирака в частотной области, т.е. одна синусоида, соответствует сигналу постоянной величины во временной области. Упрощенно и кратко, сигнал проявляет меньше изменений в одной области, чем в другой. Следовательно, при использовании предсказания или дельта-кодирования повышается эффективность кодирования, если огибающая спектра кодируется как во временном направлении, так и в частотном направлении в зависимости от характеристик сигнала.

Краткое описание чертежей

Настоящее изобретение ниже описывается посредством иллюстративных примеров, не ограничивающих объем или сущность изобретения, с ссылкой на прилагаемые чертежи, на которых

на фиг.1а-1b изображена равномерная и соответственно неравномерная дискретизация по времени огибающей спектра;

на фиг.2а-2b определено и изображено использование четырех классов гранул;

на фиг.3а-3b представлены два примера гранул и соответствующие сигналы управления;

на фиг.4а-4с представлена система передачи сигналами положения;

на фиг.5 изображено дельта-кодирование с переключением время/частота;

на фиг.6 представлена блок-схема кодера, использующего кодирование огибающей согласно изобретению;

на фиг.7 представлена блок-схема декодера, использующего кодирование огибающей согласно изобретению.

Описание предпочтительных вариантов выполнения

Описанные ниже варианты выполнения являются просто иллюстративными в отношении принципов настоящего изобретения для эффективного кодирования огибающей. Понятно, что для специалиста в этой области техники очевидны модификации и изменения устройств и описанных в них деталей. В действительности заявленное изобретение ограничивается только объемом нижеследующей формулы изобретения, а не конкретными деталями, представленными описанием и объяснением вариантов выполнения.

Формирование данных огибающей

Общим для большинства кодеров

аудиосигналов и речи является то, что как данные огибающей, так и данные остатка передаются и объединяются во время синтеза в декодере. Двумя исключениями являются кодеры, использующие перцептуальное замещение шумом ("Improving Audio Codes by Noise Substitution", D. Schultz, JAES, v.44, №7/8, 1996), и кодеры, использующие дублирование полосы спектра. В случае дублирования полосы спектра, рассматривая полосу высоких частот, должна передаваться только грубая структура спектра, так как сигнал остатка восстанавливается из полосы низких частот. Это накладывает более высокие требования на то, как сформировать данные огибающей, в частности, вследствие отсутствия "времязадающей" информации, содержащейся в первоначальном сигнале остатка. Эта проблема ниже будет продемонстрирована посредством примера.

На фиг.1 изображено частотно-временное представление музыкального сигнала, в котором длительные аккорды объединяются с крутыми переходными процессами, имеющими в основном высокочастотное содержимое. В полосе низких частот аккорды имеют высокую мощность, а мощность переходных процессов низкая, тогда как в полосе высоких частот справедливо противоположное. В данных огибающей, которые формируются в течение временных интервалов, когда присутствуют переходные процессы, преобладает высокая мощность нестационарных переходных процессов. В процессе дублирования полосы пропускания в декодере огибающая спектра транспонированного сигнала оценивается с использованием этого же мгновенного разрешения по времени и по частоте, что и используемого для анализа первоначальной полосы высоких частот. Затем выполняется выравнивание транспонированного сигнала на основе расхождения в огибающих спектра. Например, вычисляются коэффициенты усиления корректирующего огибающую банка фильтров как квадратный корень отношения средней мощности первоначального сигнала и транспонированного сигнала. Для этого типа сигнала возникает проблема: транспонированный сигнал имеет такое же отношение мощности "аккорд-переходный процесс", что и в полосе низких частот. Коэффициенты усиления, необходимые для коррекции транспонированных переходных процессов до корректного уровня, вызывают, таким образом, усиление транспонированных аккордов относительно первоначального уровня в полосе высоких частот в течение всей длительности данных огибающей, включающих энергию переходного процесса. Эти моментально слишком громкие фрагменты аккордов воспринимаются как предыдущее эхо и последующее эхо для переходного процесса, см. фиг.1а. Этот тип искажений ниже называется предыдущим и последующим эхом, вносимым коэффициентом усиления. Это явление может быть устранено постоянным обновлением данных огибающей с такой высокой скоростью, чтобы гарантировать, что промежуток времени между обновлением и произвольно расположенным переходным процессом был достаточно малым, чтобы он не был разрешимым для слуха человека. Это

решение, однако, чрезмерно увеличивает количество передаваемых данных и, таким образом, является недопустимым.

Поэтому представлена новая схема формирования данных огибающей. Решением является сохранение низкой скорости обновления во время тональных пассажей, которые составляют основную часть типичной программы, и посредством детектора переходных процессов локализация положения переходных процессов и обновление данных огибающей вблизи переднего фронта, см. фиг.1b. Это устраняет предыдущее эхо, вносимое коэффициентом усиления. Для того чтобы хорошо представить затухание переходного процесса, скорость обновления моментально увеличивается во временном интервале после начала переходного процесса. Это устраняет последующее эхо, вносимое коэффициентом усиления. Временное сегментирование во время затухания не является настолько критичным, насколько является определение начала переходного процесса, как будет объяснено ниже. Для того чтобы компенсировать меньшие временные шаги, могут быть использованы большие шаги частоты во время переходного процесса при сохранении в заданных пределах размера данных. Неравномерная дискретизация по времени и по частоте, как вкратце описано выше, применима к кодированию огибающей на основе как банка фильтров, так и на линейном предсказании. Различный порядок предсказания может быть использован для переходных и квазистационарных (тональных) сегментов.

В случае основанных на предсказании кодеров в уровне техники не известны тщательно разработанные схемы переключения разрешения по времени и по частоте. Однако некоторые кодеры на основе банков фильтров используют переменное разрешение по времени и по частоте. Обычно это достигается переключением размера банка фильтров. Такое изменение размера не может произойти мгновенно, необходимы так называемые окна перехода, и, таким образом, точки обновления не могут быть выбраны произвольно. При использовании дублирования полосы спектра или любого другого способа высокочастотного восстановления задача другая - банк фильтров может быть выполнен так, чтобы удовлетворять необходимому наивысшему разрешению как по времени, так и по частоте для выделения соответствующих представлений огибающей. Таким образом, неравномерная дискретизация по времени и по частоте огибающей спектра может быть получена посредством адаптивного группирования выборок поддиапазонов от банка фильтров фиксированного размера в "полосы частот" и "временные сегменты". Затем вычисляется одна выборка огибающей на полосу и сегмент. Ниже в описании "разрешение по частоте" относится к конкретному набору полос частоты, коэффициентам кодирования с линейным предсказанием или т.д., используемым при оценке огибающей для конкретного временного сегмента. Другими словами, с точки зрения кодирования огибающей высокое разрешение по частоте или высокое разрешение по времени может быть получено

мгновенно.

С синтаксической точки зрения все практические битовые потоки кодеров содержат периоды данных, каждый из которых соответствует короткому временному сегменту входного сигнала. Временной сегмент, связанный с таким периодом данных, ниже называется "гранулой". Обычные кодеры используют гранулы фиксированной длины. Наличие границ гранулы налагает ограничения на схему временных сегментов, используемых для оценки огибающей. Алгоритм, который формирует эти временные сегменты, может указать, что "граница" сегмента необходима в определенном положении и что последующий сегмент должен иметь определенную длину. Однако, если граница гранулы попадает внутрь этого интервала вследствие гранул фиксированной длины, сегмент должен быть разделен на две части. Это имеет два последствия. Во-первых, увеличивается количество сегментов для кодирования, возможно, увеличивая количество данных для передачи. Во-вторых, принудительные границы могут формировать сегменты, которые слишком короткие для надежных оценок средней мощности. Для того чтобы устранить эти недостатки, в настоящем изобретении используются гранулы переменной длины. Это требует просмотра вперед в кодере, а также дополнительной буферизации в декодере.

Пусть термин "сетка" обозначает временные сегменты и соответствующие разрешения по частоте для использования с конкретным сигналом, и "локальная сетка" обозначает сетку одной гранулы. Ясно, что сетка должна быть передана сигналами на декодер для правильного декодирования выборок огибающей. Однако в приложениях с низкой скоростью передачи битов количество битов для такого "сигнала управления" должно сохраняться минимальным. В настоящем изобретении предлагаются две схемы передачи сигналов. Перед их подробным описанием установим "систему исходных понятий" и некоторые критерии разработки.

Пусть шаг квантования по времени для огибающей спектра будет равен  $T_q$ . Эти шаги можно рассматривать как "субгранулы", которые сгруппированы в вышеупомянутые временные сегменты. В общем случае, гранула включает в себя  $S$  субгранул, где  $S$  меняется от гранулы к грануле. Количество возможных комбинаций сегментов в грануле, изменяемое от одного сегмента для всей гранулы до  $S$  сегментов, определяется

$$C = \sum_{n=0}^S \binom{S}{n} = 2^S \quad (1)$$

Для того чтобы передать сигналами  $C$  состояний, потребуется наименьшее целое наим цел  $(\ln_2(C)) = \text{наим цел}(\ln_2(2^S)) = S$  битов, соответствующих одному биту на субгранулу. Произвольное разбиение гранулы может быть передано сигналами посредством  $S-1$  битов, представляющих последовательные субгранулы, с указанием, присутствует ли передняя граница сегмента на соответствующей субгрануле или нет. (В этом случае не нужно передавать сигналами границы первой и последней гранулы.) Так как  $S$  переменная, она должна передаваться

сигналами, и, если эта схема объединяется с кодеком полосы низких частот с гранулами фиксированной длины, также должно передаваться сигналами положение относительно гранул постоянной длины. Разрешение по частоте сегментов может быть передано сигналами с динамически распределяемыми битами управления, например, один бит на сегмент. Ясно, что такой непосредственный способ может привести к недопустимо высокому количеству битов сигнала управления.

Как будет показано ниже, многие из состояний, описываемых уравнением (1), не являются равновероятными и также формируют слишком большое количество данных огибающей для практического использования при ограниченной скорости передачи битов.

Минимальный временной диапазон между последовательными переходными процессами в музыкальной программе может быть оценен следующим образом. В нотной записи "ритм" описывается тактовым размером, выраженным как дробь A/B, где A обозначает число "долей" на такт, и 1/B - тип ноты, соответствующей одной доле, например 1/4 ноты, обычно называемая четвертной нотой. Пусть t обозначает темп в долях в минуту (доля/мин). Время на ноту типа 1/C тогда определяется следующим образом:

$$T_n = (60/t) \cdot (B/C), \text{ с} \quad (2)$$

Большинство музыкальных пьес попадают в диапазон 70-160 доля/мин, и в тактовом размере 4/4 самые быстрые ритмические модели для большинства практических случаев состоят из нот 1/32 или из 32 нот. Это дает минимальное время  $T_{nmin} = (60/160) \cdot (4/32) = 47$  мс. Конечно, могут иметь место более короткие временные периоды, чем этот, но такие быстрые последовательности (>21 событий в секунду) дают почти характер гудения, и он не должен быть полностью различимым.

Также должно быть установлено необходимое разрешение во времени  $T_q$ . В некоторых случаях у переходного сигнала его подлежащая восстановлению основная энергия находится в полосе высоких частот. Это означает, что кодированная огибающая спектра должна включать всю "временную" информацию. Требуемая точность во времени определяет, таким образом, разрешение, необходимое для кодирования передних фронтов.  $T_q$  намного меньше минимального периода  $T_{nmin}$  ноты, так как ясно могут быть услышаны небольшие временные отклонения в пределах периода. Во многих случаях, однако, переходный процесс имеет значительную энергию в полосе низких частот. Вышеописанное предыдущее эхо, вносимое коэффициентом усиления, должно попадать в пределы так называемого времени  $T_m$  предмаскирования или маскирования "назад" слуховой системы человека, чтобы быть неслышимым.  $T_q$ , следовательно, должен удовлетворять двум условиям

$$T_q << T_{nmin}, \quad (3)$$

$$T_q < T_m. \quad (4)$$

Очевидно, что  $T_m < T_{nmin}$  (иначе ноты будут такие быстрые, что их нельзя будет

разобрать), и согласно ("Modeling the Additivity of Nonsimultaneous Masking", Hearing Res., v. 80, p. 105-118 (1994))  $T_m$  составляет до 10-20 мс. Так как  $T_{nmin}$  находится в диапазоне 50 мс, обоснованный выбор  $T_q$  согласно уравнению (3) приводит к тому, что также удовлетворяется второе условие. При выборе  $T_q$  необходимо также учитывать, конечно, точность обнаружения переходного процесса в кодере и разрешение во времени в банке фильтров анализа/синтеза.

Отслеживание задних фронтов менее критично по нескольким причинам. Во-первых, положение завершения ноты имеет малое или не имеет влияния на воспринимаемый ритм. Во-вторых, у большинства инструментов не проявляются крутые задние фронты, а скорее плавная кривая затухания, т.е. не существует четко определенного времени завершения ноты. В-третьих, время постмаскирования или маскирования "вперед" значительно больше, чем время предмаскирования.

Итак, суммируя вышеизложенное, могут быть сделаны следующие упрощения без потери или с малой потерей качества для практических сигналов:

1) необходимо передавать только начальное положение переходного процесса с самой высокой точностью  $T_q$ ;

2) в данных огибающей необходимо полностью различать только переходные процессы, разделенные  $T_p \gg T_q$ .

Для того чтобы уменьшить количество служебных сигналов, обе системы согласно настоящему изобретению используют два режима дискретизации по времени: равномерная и неравномерная дискретизация по времени. Равномерный режим используется во время квазистационарных пассажей, в соответствии с чем используются сегменты фиксированной длины, и требуется передача нескольких дополнительных сигналов. Вблизи переходных процессов система переключается в неравномерный режим работы и используются гранулы переменной длины, приводящие к соответствию с идеальной общей сеткой.

Система передачи сигналами класса В первой системе гранулы делятся на четыре класса, и сигналы управления адаптируются к конкретным требованиям каждого класса. Классы определены на фиг.2а. Класс "ФиксФикс" соответствует обычным гранулам постоянной длины. Класс "ФиксПер" имеет подвижную границу останова, которая позволяет производить изменение длины гранулы. Класс "ПерФикс" имеет подвижную границу начала, тогда как граница останова фиксированная. Последний класс "ПерПер" имеет переменные границы на обоих концах. Все переменные границы могут смещаться на  $-a/+b$  относительно "номинальных положений".

На фиг.2b приведен пример последовательности гранул. По умолчанию в системе установлен класс "ФиксФикс". Детектор переходных процессов (или психоакустическая модель) работает в области времени перед текущей гранулой, как изображено на фигуре. Когда обнаруживается переходный процесс, используется гранула класса "ФиксПер" - система переключается из

равномерного режима в неравномерный режим работы. Обычно за этой гранулой следует гранула класса "ПерФикс", так как в течение большей части времени переходные процессы разделены рядом гранул для всех практических вариантов длины гранул. В случае переходных процессов в последовательных фреймах могут быть использованы фреймы класса "ПерПер".

На фиг.3а приведен пример пары классов "ФиксПер - ПерФикс" и соответствующий сигнал управления. Присутствует один переходный процесс, и передний фронт (квантованный  $T_q$ ) обозначен  $t$ . Первой частью битового потока является сигнал "класса". Так как используются четыре бита. В случае классов "ФиксПер" или "ПерФикс" следующий сигнал определяет положение переменной границы, выраженной как смещение от номинального положения. Эта граница называется "абсолютной границей". Границы сегментов в гранулах определяются при помощи "относительных границ". Абсолютная граница используется как исходная точка, а другие границы определяются как итоговые расстояния от исходной точки. Количество относительных границ переменное и передается сигналами на декодер после абсолютной границы. Нулевое число означает, что гранула включает только один временной сегмент. Таким образом, в случае класса "ФиксПер" длины сегментов передаются сигналами в обратной последовательности, двигаясь от абсолютной границы на конце гранулы. Длина первого сегмента в грануле "ФиксПер" выводится из относительных границ и суммарной длины и не передается сигналами. Сигналы относительных границ класса "ПерФикс" вводятся в битовый поток в прямой последовательности, посредством чего длина последнего сегмента исключается. Порядок сигналов битового потока идентичен порядку класса "ФиксПер", т.е. [класс, абсолютная граница, число относительных границ, относительная граница 0, относительная граница 1, ... относительная граница N-1]. На фигуре сигналы приведены "открытым текстом" вместо фактических слов двоичного кода, посылаемых в битовом потоке.

На фиг.3б изображено альтернативное кодирование сигнала. Переменная граница обеспечивает универсальность при группировании сегментов при данной общей сетке. На этом уровне, таким образом, может быть выполнено некоторое управление полезной нагрузкой, например выравнивание числа битов на гранулу. Это может упростить работу кодера полосы низких частот. После достаточного предварительного просмотра может быть выполнено многопроходное кодирование и использована оптимальная комбинация локальных сеток.

Для того чтобы уменьшить набор символов для передачи сигналами относительных границ и, тем самым, количество битов на символ, эти длины могут квантоваться целым числом, кратным ( $>1$ )  $T_q$ , если абсолютная граница имеет точность  $T_q$ . В этом случае абсолютная граница, в дополнение к вышеописанным функциям, служит для выравнивания группы границ вокруг переходного процесса с точностью  $T_q$ . Другими словами, наивысшая точность

всегда доступна для кодирования передних фронтов переходных процессов, а грубое разрешение используется для отслеживания затухания. В фреймах класса "ПерПер" используется комбинация передачи сигналами "ФиксПер" и "ПерФикс", например чередующаяся: [класс, абсолютная левая граница, то же самое для правой, число относительных левых границ, то же самое для правой, [относительная левая граница 0, ... относительная левая граница N-1], [то же самое для правой]]. Этот класс предлагает самую большую гибкость при выборе локальной сетки за счет увеличенного количества служебных сигналов. Наконец, класс "ФиксФикс" не требует других сигналов, чем сигнал класса сам по себе, причем в этом случае, например, используются два сегмента (одинаковой длины). Однако допустимо добавить сигнал, который позволяет производить выбор из набора заданных сеток. Например, огибающая спектра может быть вычислена для двух сегментов, и если две огибающие не отличаются более чем на некоторую величину, то посылаются только один набор данных огибающей.

До сих пор описывалось только сегментирование во времени. По многим причинам может быть желательным передавать сигналами на декодер то, какая из границ соответствует переднему фронту переходного процесса. Это может быть осуществлено посылкой "указателя", который указывает на соответствующую границу. Направление исходной точки может повторять направление относительных границ, и нулевая величина означает, что начало переходного процесса не присутствует в текущей грануле. Кроме того, также должно быть определено разрешение по частоте (число оценок мощности или порядок предсказания), используемое для индивидуальных сегментов. Оно может быть явно передано сигналами как в "системе исходных понятий" или неявно, т.е. разрешение связано с длинами сегментов и, возможно, положением указателя.

При использовании подверженных ошибкам каналов передачи важно избегать распространения ошибок. В вышеописанной системе локальная сетка полностью описывается сигналом управления соответствующей гранулы. Следовательно, в сигнале управления нет зависимостей между фреймами. Это означает то, что границы гранул "перекодируются", так как пересечение гранул передаются сигналами в обеих последовательных гранулах. Эта избыточность может быть использована для обнаружения простых ошибок - если границы не соответствуют, произошла ошибка передачи и может быть активизировано маскирование ошибок.

Система передачи сигналами положения. Вторая система, ниже называемая "системой передачи сигналами положения", предназначена для применений с очень низкой скоростью передачи битов. Ранее установленные правила разработки используются в еще большей степени, чтобы еще больше уменьшить количество битов сигнала управления. Согласно настоящему изобретению информация о начале переходного процесса может быть использована для неявной передачи

сигналами границ сегментов и разрешений по частоте вблизи переходных процессов. Ниже это описывается с предположением, что номинальный размер гранул равен  $N$  субгранулам, выбранным согласно  $NT_q \ll T_{nmin}$ , т.е. максимум один переходный процесс, вероятно, происходит в пределах гранулы, см. фиг.4а, где  $N=8$ . Как показано на фиг.4b, используется детектор переходных процессов, работающий на интервалах длиной  $N$ , расположенных на расстоянии  $N/2$  перед текущей гранулой. Когда обнаруживается переходный процесс, устанавливается флаг, связанный с этой областью. В данном примере детектор переходных процессов обнаружил переходный процесс в субгрануле 2 в момент времени  $n-1$  и переходный процесс в субгрануле 3 в момент времени  $n$ . Эти положения пол( $n-1$ ) и пол( $n$ ), а также соответствующие флаги флаг( $n-1$ ) и флаг( $n$ ) используются в качестве входных величин алгоритма формирования сетки, и соответствующая локальная сетка для гранулы  $n$  может быть такой, как показанная на фиг.4с. Как показано на фигуре, субгранула 3 гранулы в момент времени  $n-1$  включена в сетку по времени и по частоте гранулы  $n$ . Единственными сигналами, передаваемыми в битовом потоке, являются флаг( $n$ ) [1 бит] и пол( $n$ ) [наим цел( $\ln_2(N)$ ) бит]. Алгоритм сетки также известен декодеру, следовательно, эти сигналы, вместе с соответствующими сигналами предыдущей гранулы  $n-1$ , достаточны для однозначного восстановления сетки, используемой декодером. Если переходный процесс не обнаружен, то сигнал положения является устаревшим и может быть заменен, например, 1-битовым сигналом, указывающим, что используется один или два сегмента. Таким образом, работа в равномерном режиме идентична работе системы передачи сигналами класса.

Эта система может рассматриваться как конечный автомат, в котором вышеописанные сигналы управляют переходами из состояния в состояние, и состояния определяют локальные сетки. Ясно, что состояния могут быть представлены таблицами, хранимыми как в кодере, так и в декодере. Так как сетки являются жестко запрограммированными, то нет возможности адаптивно изменять полезную нагрузку. Приемлемым решением является сохранение приблизительно постоянным размера матрицы данных по времени и по частоте (например, числа оценок мощности). Предполагая, что число масштабных множителей или коэффициентов в сегменте с высоким разрешением в два раза больше, чем в сегменте с низким разрешением, то один сегмент с высоким разрешением может быть заменен на два сегмента с низким разрешением.

Кодирование масштабных множителей с переключением время/частота

При использовании частотно-временного преобразования можно показать, что импульс во временной области соответствует равномерному спектру в частотной области, и "импульс" в частотной области, т.е. одна синусоида, соответствует квазистационарному сигналу во временной области. Другими словами, сигнал обычно

проявляет больше свойств переходных процессов в одной области, чем в другой. На спектрограмме, т.е. частотно-временном матричном отображении, это свойство очевидно и может быть выгодно использовано при кодировании огибающих спектра.

Тональный стационарный сигнал может иметь очень разреженный спектр, не пригодный для дельта-кодирования в частотном направлении, но хорошо подходящий для дельта-кодирования во временном направлении, и наоборот. Это изображено на фиг.5. В нижеследующем описании вектор масштабных множителей, вычисленный в момент времени  $n_0$ , представляет огибающую спектра

$$Y(k, n_0) = [a_1, a_2, a_3 \dots a_k \dots a_N], \quad (5)$$

где  $a_1 \dots a_N$  - величины амплитуды для различных частот. Обычно кодируют разность между соседними величинами в частотном направлении в данный момент времени, что дает

$$D(k, n_0) = [a_2 - a_1, a_3 - a_2, \dots a_N - a_{(N-1)}]. \quad (6)$$

Для возможности декодирования этого необходимо передавать начальную величину  $a_1$ . Как указано выше, эта схема дельта-кодирования может оказаться наименее эффективной, если спектр содержит только несколько установившихся нот. Результатом этого может быть высокая скорость передачи битов, чем обычное ИКМ-кодирование. Для решения этой проблемы предлагается способ с переключением время/частота, ниже называемый частотно-временным кодированием. Масштабные множители квантуются и кодируются как во временном, так и в частотном направлении. Для обоих случаев необходимое количество битов вычисляется для данной ошибки кодирования или ошибка вычисляется для данного количества битов. На основе этого выбирается наиболее выгодное направление кодирования.

В качестве примера могут быть использованы избыточное ДИКМ-кодирование (кодирование с дифференциальной импульсно-кодовой модуляцией) и избыточное кодирование Хаффмана. Вычисляются два вектора  $D_f$  и  $D_t$

$$D_f(k, n_0) = [a_2 - a_1, a_3 - a_2, \dots a_N - a_{(N-1)}], \quad (7)$$

$$D_t(k, n_0) = [a_1(n_0) - a_1(n_0 - 1), a_2(n_0) - a_2(n_0 - 1) \dots a_N(n_0) - a_N(n_0 - 1)]. \quad (8)$$

Соответствующие таблицы Хаффмана - одна для частотного направления и одна для временного направления - указывают число битов, необходимое для кодирования векторов. Кодированный вектор, требующий наименьшее число битов для кодирования, представляет предпочтительное направление кодирования. Таблицы первоначально могут быть сгенерированы с использованием некоторого минимального расстояния в качестве критерия переключения время/частота.

Начальные величины передаются всякий раз, когда огибающая спектра кодируется в частотном направлении, но не тогда, когда кодируется во временном направлении, так

как они имеются в наличии у декодера от предыдущей огибающей. Предлагаемый алгоритм также требует передачи дополнительной информации, а именно флага время/частота, указывающего, в каком направлении кодировалась огибающая спектра. Частотно-временной алгоритм может быть выгодно использован с несколькими другими схемами кодирования представления огибающей с масштабными множителями, кроме кодирования ДИКМ и Хаффмана, такими как АДИКМ-кодирования (кодирование с адаптивной дифференциальной импульсно-кодовой модуляцией), кодирование с линейным предсказанием и векторное квантование. Предлагаемый частотно-временной алгоритм создает значительное снижение скорости передачи битов для данных огибающей спектра.

Практическое выполнение

Пример кодера согласно изобретению приведен на фиг.6. Входной аналоговый сигнал подается на аналого-цифровой преобразователь 601, формирующий цифровой сигнал. Цифровой аудиосигнал подается на перцептуальный кодер 602 аудиосигналов, в котором выполняется исходное кодирование. Кроме того, цифровой сигнал подается в детектор 603 переходных процессов и в банк 604 фильтров анализа, который разделяет сигнал на его спектральные эквиваленты (сигналы поддиапазонов). Детектор переходных процессов может воздействовать на сигналы поддиапазонов от банка анализа, но для универсальности здесь предполагается воздействие непосредственно на цифровые выборки временной области. Детектор переходных процессов разделяет сигнал на гранулы и определяет согласно изобретению, отмечены ли флагом как переходный процесс субгранулы в пределах гранул. Эта информация посылается на блок 605 группирования огибающей, который определяет частотно-временную сетку, которую необходимо использовать для текущей гранулы. Согласно сетке блок объединяет равномерно дискретизированные сигналы поддиапазонов, формируя неравномерно дискретизированные величины огибающей. В качестве примера, эти величины могут представлять среднюю плотность мощности сгруппированных выборок поддиапазонов. Величины огибающей, вместе с информацией группирования, подаются в блок 606 кодера огибающей. Этот блок определяет, в каком направлении (по времени или по частоте) кодировать величины огибающей. Результирующие сигналы, выходной сигнал кодера аудиосигналов, информация широкополосной огибающей и сигналы управления подаются в мультиплексор 607, формирующий последовательный битовый поток, который передается или хранится.

Декодер согласно изобретению изображен на фиг.7, использующий транспонирование дублирования полосы спектра в качестве примера формирования отсутствующего сигнала остатка. Демультимплексор 701 восстанавливает сигналы и подает соответствующую часть на декодер 702 аудиосигналов, который создает цифровой аудиосигнал полосы низких частот. Информация об огибающей подается с

демультимплексора в блок 703 декодирования огибающей, который посредством использования данных управления определяет, в каком направлении кодируется текущая огибающая, и декодирует данные. Сигнал полосы низких частот от декодера аудиосигналов направляется в модуль 704 транспонирования, который формирует дублируемый сигнал полосы высоких частот из полосы низких частот. Сигнал полосы высоких частот подается в банк 706 фильтров анализа, который такого же типа, что и в кодере. Сигналы поддиапазонов объединяются в блоке 707 группирования с масштабными множителями. Посредством использования данных управления от демультимплексора принимается этот же тип объединения и частотно-временного распределения выборок поддиапазонов, что и в кодере. Информация об огибающей от демультимплексора и информация от блока группирования с масштабными множителями обрабатывается в модуле 708 управления коэффициентом усиления. Модуль вычисляет коэффициенты усиления, которые должны быть применимы к выборкам поддиапазонов перед воссоединением в блоке 709 банка фильтров синтеза. Выходным сигналом банка фильтров синтеза является, таким образом, аудиосигнал полосы высоких частот с откорректированной огибающей. Этот сигнал добавляется к выходному сигналу блока 705 задержки, который подается с аудиосигналом полосы низких частот. Задержка компенсирует время обработки сигнала полосы высоких частот. В заключение, полученный цифровой широкополосный сигнал преобразуется в аналоговый аудиосигнал в цифроаналоговом преобразователе 710.

### Формула изобретения:

1. Способ кодирования огибающей спектра для входного сигнала, занимающего некоторую полосу частот, причем полоса частот включает в себя области определенных частот, при этом входной сигнал представлен его исходно-кодированной версией, занимающей некоторую полосу частот, которая не включает в себя области определенных частот, причем огибающая спектра входного сигнала в областях определенных частот представлена с помощью грубого представления огибающей спектра и точного представления огибающей спектра, при этом точное представление огибающей спектра представляет собой сигнал остатка, заключающийся в том, что выполняют (603) статистический анализ входного сигнала, на основе результата статистического анализа формируют (604, 605, 606) данные по грубому представлению огибающей спектра для областей определенных частот с помощью дискретизации огибающей спектра в областях определенных частот с переменным разрешением по времени или с переменным разрешением по частоте, причем разрешение по времени или разрешение по частоте, выбранное для некоторого момента времени, зависит от результата статистического анализа входного сигнала в этот момент времени, формируют сигнал управления, описывающий переменное разрешение по времени или переменное разрешение по частоте, и формируют (607) кодированный

входной сигнал с помощью мультиплексирования исходно-кодированной версии, данных по грубому представлению огибающей спектра и сигнала управления, причем кодированный входной сигнал не включает в себя сигнал остатка.

2. Способ по п.1, отличающийся тем, что при формировании (604, 605, 606) данных по грубому представлению огибающей спектра для областей определенных частот выбирают сетку разрешения по времени/по частоте, используемую для грубого представления огибающей спектра, при этом сигнал управления генерируют для описания упомянутой сетки.

3. Способ по п.1 или 2, отличающийся тем, что при формировании данных по грубому представлению огибающей спектра получают элементы частотного/временного представления входного сигнала, группируют элементы частотного/временного представления входного сигнала и вычисляют масштабный множитель для каждой группы.

4. Способ по п.3, отличающийся тем, что при операции получения используют банк фильтров.

5. Способ по п.4, отличающийся тем, что банк фильтров имеет фиксированный размер.

6. Способ по п.1, отличающийся тем, что при формировании данных по грубому представлению огибающей спектра для областей определенных частот используют устройство линейного предсказания.

7. Способ по п.1, отличающийся тем, что при выполнении статистического анализа используют детектор переходных процессов.

8. Способ по п.1, отличающийся тем, что при формировании данных по грубому представлению огибающей спектра переключают мгновенное разрешение из комбинации по умолчанию с более высоким разрешением по частоте и более низким разрешением по времени в комбинацию с более низким разрешением по частоте и более высоким разрешением по времени в начале переходного периода для получения переменного разрешения по времени переменного разрешения по частоте.

9. Способ по п.1, отличающийся тем, что операция формирования сигнала управления предназначена для формирования такого сигнала управления, который описывает положения в пределах гранулы постоянной скорости обновления, при этом операция выполнения статистического анализа предназначена для получения постоянной скорости обновления и операция формирования (604, 605, 606) данных по грубому представлению огибающей спектра предназначена для выбора мгновенного разрешения на основе положений переходных процессов во входных сигналах в пределах текущей и соседних гранул посредством использования правил, доступных кодеру и декодеру.

10. Способ по п.9, отличающийся тем, что операция формирования сигнала управления предназначена для формирования такого сигнала управления, что передают сигналами не более одного положения на гранулу.

11. Способ по п.1, отличающийся тем, что операция формирования (604, 605, 606) данных по грубому представлению огибающей спектра предназначена для использования гранул переменной длины.

12. Способ по п.11, отличающийся тем, что используют четыре класса гранул, в соответствии с чем первый класс имеет границы гранул с фиксированным положением и длину  $L$ , второй класс имеет границу начала с фиксированным положением и границу останова с переменным положением, третий класс имеет границу начала с переменным положением и границу останова с фиксированным положением, четвертый класс имеет границы начала и останова с переменным положением и фиксированные положения совпадают с положениями исходной точки, разделенными упомянутым расстоянием  $L$ , а переменные положения могут быть смещены  $[-a, b]$  относительно положений исходной точки.

13. Способ по п.3, отличающийся тем, что при формировании (604, 605, 606) данных по грубому представлению огибающей спектра кодируют масштабные множители как во временном направлении, так и в частотном направлении, причем в каждый момент времени определяют наиболее выгодное направление и при этом наиболее выгодное направление выбирают при упомянутой операции кодирования.

14. Способ по п.13, отличающийся тем, что при формировании (604, 605, 606) данных по грубому представлению огибающей спектра кодируют масштабные множители как во временном направлении, так и в частотном направлении, причем для операции кодирования выбирают направление, которое создает наименьшую ошибку кодирования для данного числа битов.

15. Способ по п.13, отличающийся тем, что при формировании (604, 605, 606) данных по грубому представлению огибающей спектра кодируют масштабные множители как во временном направлении, так и в частотном направлении, причем для операции кодирования выбирают направление, которое создает наименьшее количество битов для данной ошибки кодирования.

16. Способ по пп.13, 14 или 15, отличающийся тем, что при кодировании используют кодирование без потерь и используют отдельные таблицы для временного направления и частотного направления, причем результат кодирования с использованием таблиц используют для выбора направления для кодирования.

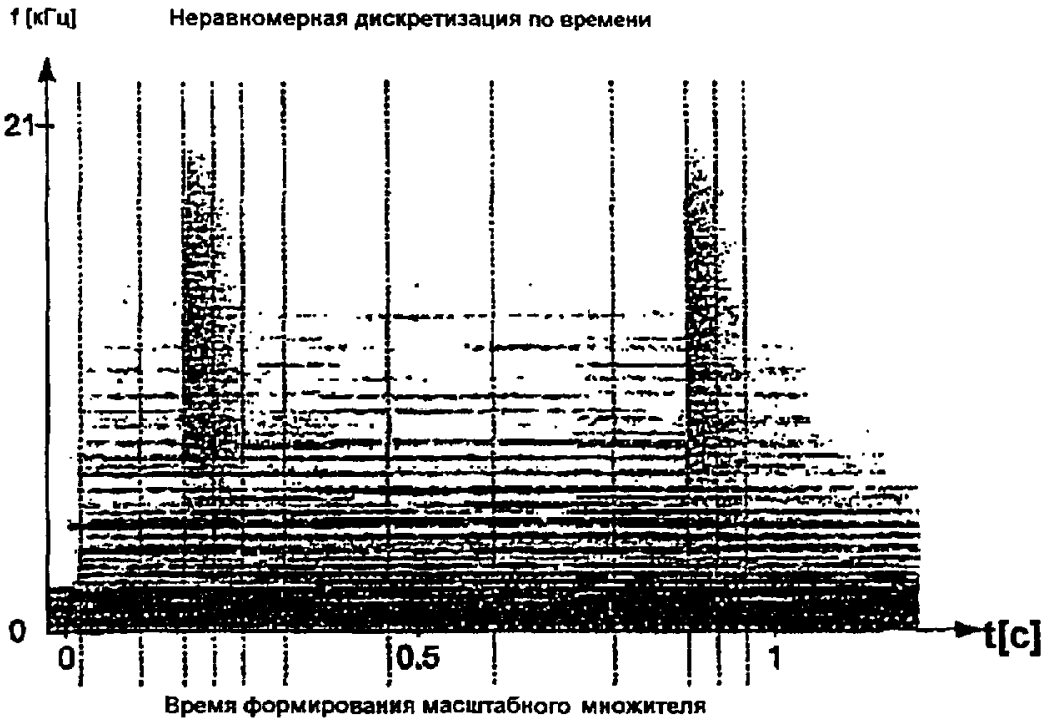
17. Устройство кодирования огибающей спектра для входного сигнала, занимающего некоторую полосу частот, причем полоса частот включает в себя области определенных частот, при этом входной сигнал представлен его исходно-кодированной версией, занимающей некоторую полосу частот, которая не включает в себя области определенных частот, причем огибающая спектра входного сигнала в областях определенных частот представлена с помощью грубого представления огибающей спектра и точного представления огибающей спектра, при этом точное представление огибающей спектра представляет собой сигнал остатка, содержащее средство (603) для выполнения статистического анализа входного сигнала, средство для формирования (604, 605, 606) данных по грубому представлению огибающей спектра для областей определенных частот на основе результата

статистического анализа с помощью выборки огибающей спектра в областях определенных частот с переменным разрешением по времени или с переменным разрешением по частоте, причем разрешение по времени или разрешению по частоте, выбранное для некоторого момента времени, зависит от результата статистического анализа входного сигнала в этот момент времени, для формирования сигнала управления, описывающего переменное разрешение по времени или переменное разрешение по частоте и для формирования (607) кодированного входного сигнала с помощью мультиплексирования исходно-кодированной версии, данных по грубому представлению огибающей спектра и сигнала управления, причем кодированный входной сигнал не включает в себя сигнал остатка.

18. Устройство декодирования огибающей спектра кодированного сигнала, включающего в себя исходно-кодированную версию оригинального сигнала, причем оригинальный сигнал занимает некоторую полосу частот, включающую в себя области определенных частот, а исходно-кодированная версия занимает некоторую полосу частот, которая не включает в себя области определенных частот, данные по грубому представлению огибающей спектра для областей определенных частот, причем данные по грубому представлению огибающей спектра представляют огибающую спектра с переменным разрешением по времени или с переменным разрешением по частоте, и сигнал управления, указывающий переменное разрешение по времени или переменное разрешение по частоте, и получения в результате из исходно-кодированного сигнала после исходного декодирования (702) декодированной версии оригинального сигнала, причем декодированная версия оригинального сигнала занимает некоторую полосу частот, которая не включает в себя области определенных частот, содержащее демультимплексор (701) для демультимплексирования кодированного сигнала для получения исходно-кодированной версии, данных по грубому представлению огибающей спектра и сигнала управления, средство (704) для формирования сигнала с дублированной полосой спектра для областей определенных частот, средство для интерпретации сигнала управления с целью определения переменного разрешения по времени или переменного разрешения по частоте, средство (708, 709) для коррекции огибающей сигнала с дублированной полосой спектра с использованием данных по грубому

представлению огибающей спектра и переменного разрешения по времени или переменного разрешения по частоте и средство для суммирования сигнала откорректированной огибающей и декодированной версии оригинального сигнала для получения декодированного сигнала, занимающего некоторую полосу частот, включающую в себя области определенных частот.

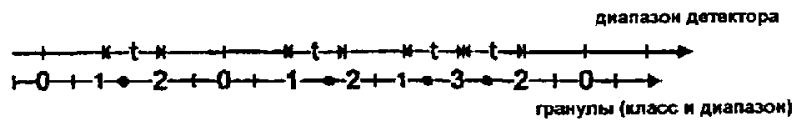
19. Способ декодирования огибающей спектра кодированного сигнала, включающего в себя исходно-кодированную версию оригинального сигнала, причем оригинальный сигнал занимает некоторую полосу частот, включающую в себя области определенных частот, а исходно-кодированная версия занимает некоторую полосу частот, которая не включает в себя области определенных частот, данные по грубому представлению огибающей спектра для областей определенных частот, причем данные по грубому представлению огибающей спектра представляют огибающую спектра с переменным разрешением по времени или с переменным разрешением по частоте, и сигнал управления, указывающий переменное разрешение по времени или переменное разрешение по частоте, и получения в результате из исходно-кодированного сигнала после исходного декодирования (702) декодированной версии оригинального сигнала, причем декодированная версия оригинального сигнала занимает некоторую полосу частот, которая не включает в себя области определенных частот, заключающийся в том, что демультимплексируют (701) кодированный сигнал для получения исходно-кодированной версии, данных по грубому представлению огибающей спектра и сигнала управления, формируют (704) сигнал с дублированной полосой спектра для областей определенных частот, интерпретируют (703) сигнал управления с целью определения переменного разрешения по времени или переменного разрешения по частоте, регулируют (708, 709) огибающую сигнала с дублированной полосой спектра с использованием данных по грубому представлению огибающей спектра и переменного разрешения по времени или переменного разрешения по частоте и суммируют сигнал откорректированной огибающей и декодированной версии оригинального сигнала для получения декодированного сигнала, занимающего некоторую полосу частот, включающую в себя области определенных частот.



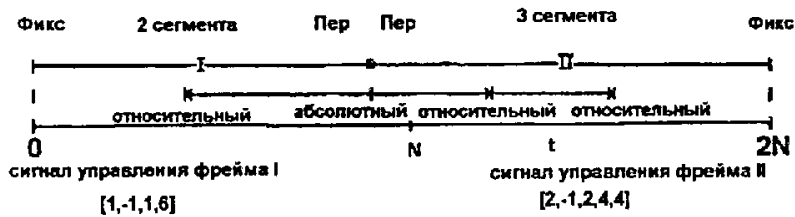
ФИГ.1b



ФИГ.2a



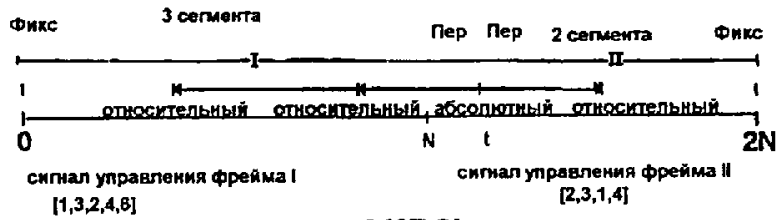
ФИГ.2b



ФИГ.3a

RU 2 2 3 6 0 4 6 C 2

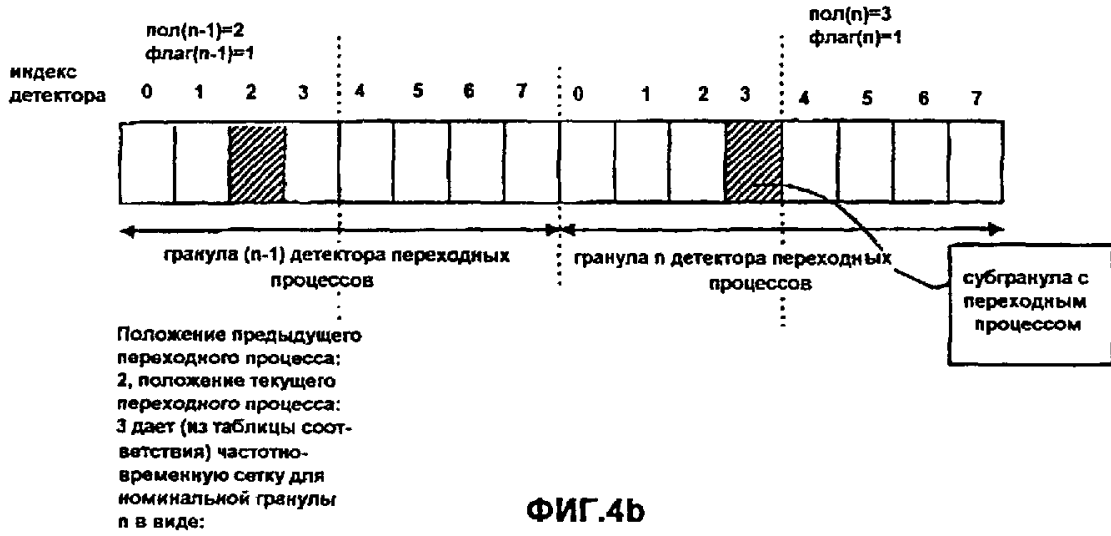
RU 2 2 3 6 0 4 6 C 2



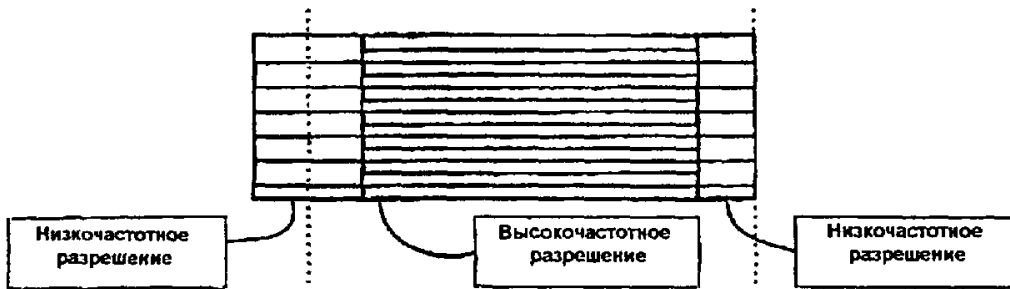
ФИГ.3б



ФИГ.4а



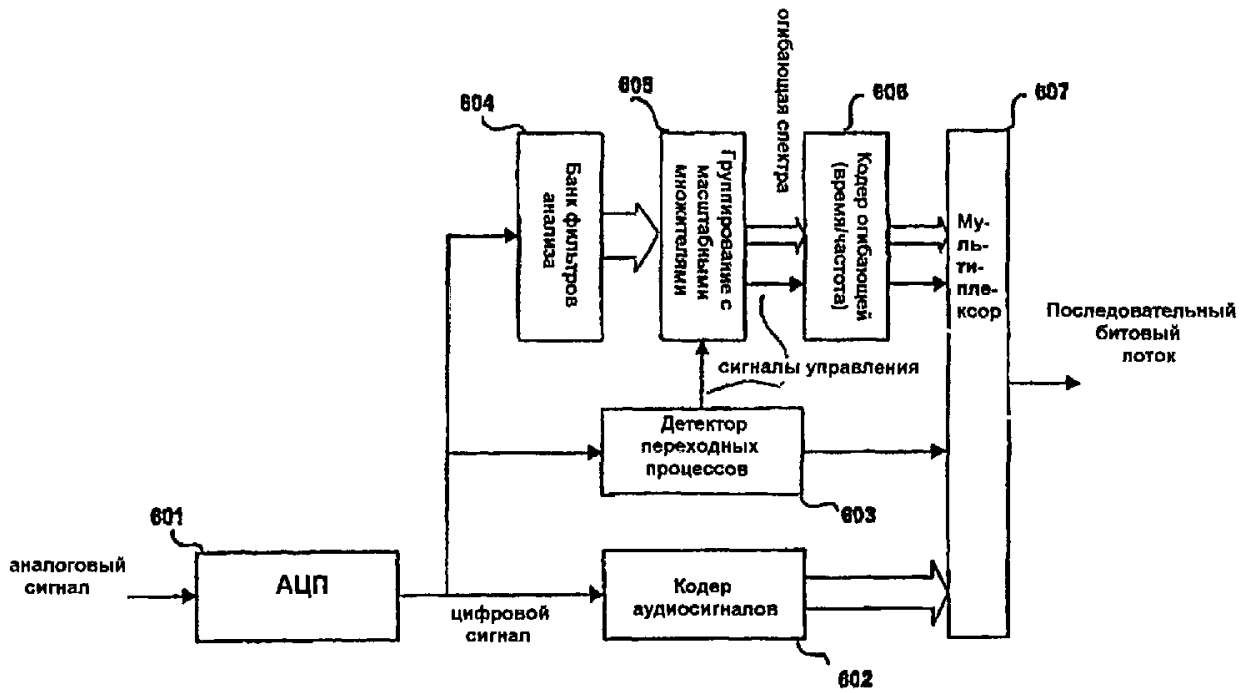
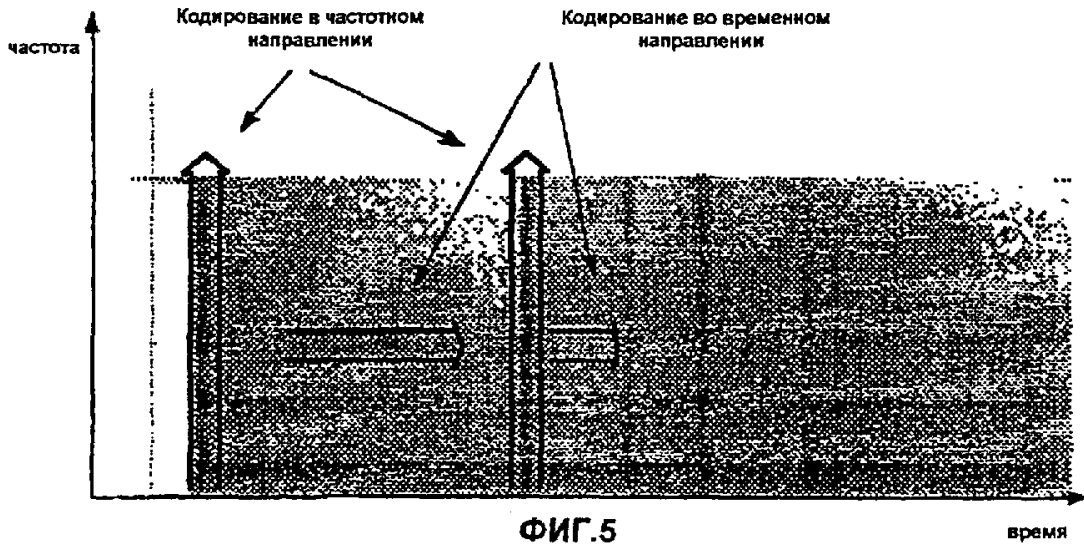
ФИГ.4б



ФИГ.4с

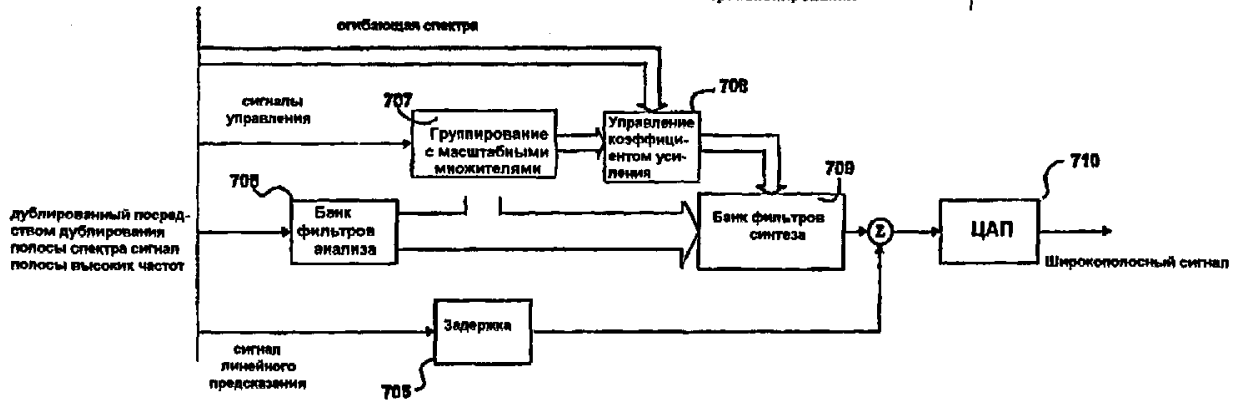
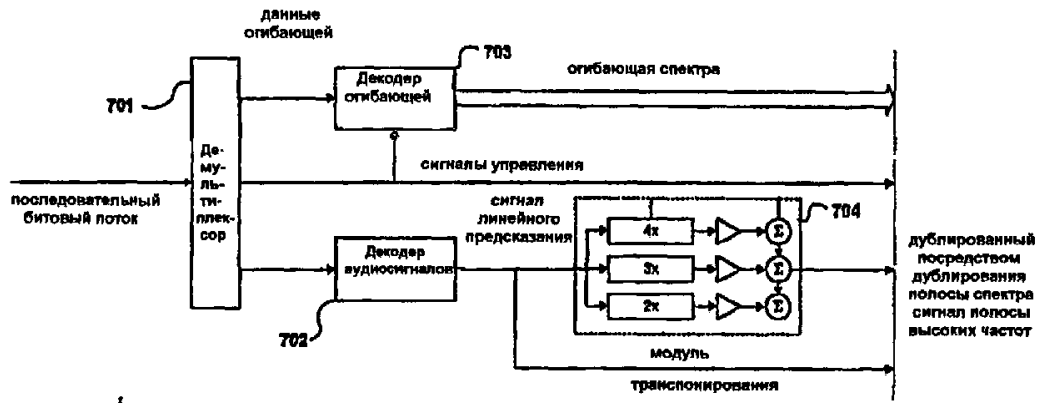
RU 2236046 C2

RU 2236046 C2



RU 2236046 C2

RU 2236046 C2



ФИГ.7

RU 2236046 C2

RU 2236046 C2