



ФЕДЕРАЛЬНАЯ СЛУЖБА
ПО ИНТЕЛЛЕКТУАЛЬНОЙ СОБСТВЕННОСТИ,
ПАТЕНТАМ И ТОВАРНЫМ ЗНАКАМ

(51) МПК

H04S 3/00 (2006.01)*G10L 19/00* (2006.01)**(12) ОПИСАНИЕ ИЗОБРЕТЕНИЯ К ПАТЕНТУ**(21), (22) Заявка: **2007136792/09, 22.02.2006**(24) Дата начала отсчета срока действия патента:
22.02.2006(30) Конвенционный приоритет:
04.03.2005 DE 102005010057.0(43) Дата публикации заявки: **10.04.2009**(45) Опубликовано: **20.12.2009** Бюл. № 35(56) Список документов, цитированных в отчете о
поиске: **WO 03/086017 A2, 16.10.2003. RU 2129336**
C1, 29.04.1999. US 6275589 B1, 14.08.2001. RU
2119259 C1, 20.09.1998. US 6023490 A,
08.02.2000.(85) Дата перевода заявки РСТ на национальную
фазу: **04.10.2007**(86) Заявка РСТ:
EP 2006/001622 (22.02.2006)(87) Публикация РСТ:
WO 2006/094635 (14.09.2006)Адрес для переписки:
129090, Москва, ул. Б.Спасская, 25, стр.3,
ООО "Юридическая фирма Городисский и
Партнеры", пат.пов. Ю.Д.Кузнецову,
рег.№ 595

(72) Автор(ы):

ПЛОГСТИС Ян (DE),
МУНДТ Харальд (DE),
ПОПП Харальд (DE)

(73) Патентообладатель(и):

ФРАУНХОФЕР-ГЕЗЕЛЛЬШАФТ ЦУР
ФЕРДЕРУНГ ДЕР АНГЕВАНДТЕН
ФОРШУНГ Е.Ф. (DE)**(54) УСТРОЙСТВО И СПОСОБ ДЛЯ ФОРМИРОВАНИЯ ЗАКОДИРОВАННОГО
СТЕРЕОСИГНАЛА АУДИОЧАСТИ ИЛИ ПОТОКА ДАННЫХ АУДИО**

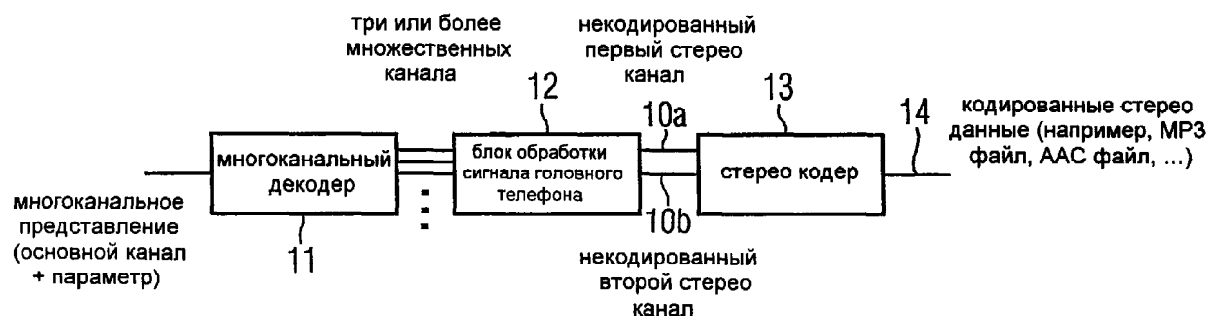
(57) Реферат:

Изобретение относится к технологии многоканального аудио и, в частности, к приложениям многоканального аудио в соединении с технологиями головных телефонов. Устройство для формирования закодированного стереосигнала из многоканального представления включает в себя многоканальный декодер (11), формирующий три или более множественных каналов из по меньшей мере одного основного

канала и параметрической информации. Упомянутые три или более множественных канала подвергаются обработке (12) сигналов головного телефона, чтобы сформировать некодированный первый стереоканал и некодированный второй канал стерео, которые затем подают на стереокодер (13), чтобы сформировать закодированный файл стерео на стороне выхода. Закодированный файл стерео может быть подан на любое подходящее устройство воспроизведения в форме

проигрывателя CD или аппаратного устройства воспроизведения так, что пользователь устройства воспроизведения не только получает нормальное впечатление стерео, но и многоканальное впечатление. Технический результат - обеспечение

эффективной концепции обработки сигналов, допускающей многоканальное качественное воспроизведение на головном телефоне на простых устройствах воспроизведения. 3 н. и 8 з.п. ф-лы, 11 ил.



Фиг. 1



FEDERAL SERVICE
FOR INTELLECTUAL PROPERTY,
PATENTS AND TRADEMARKS

(51) Int. Cl.

H04S 3/00 (2006.01)*G10L* 19/00 (2006.01)**(12) ABSTRACT OF INVENTION**(21), (22) Application: **2007136792/09, 22.02.2006**(24) Effective date for property rights:
22.02.2006(30) Priority:
04.03.2005 DE 102005010057.0(43) Application published: **10.04.2009**(45) Date of publication: **20.12.2009 Bull. 35**(85) Commencement of national phase: **04.10.2007**(86) PCT application:
EP 2006/001622 (22.02.2006)(87) PCT publication:
WO 2006/094635 (14.09.2006)

Mail address:

129090, Moskva, ul. B.Spasskaja, 25, str.3, OOO
"Juridicheskaja firma Gorodisskij i Partnery",
pat.pov. Ju.D.Kuznetsovu, reg.№ 595

(72) Inventor(s):

**PLOGSTIS Jan (DE),
MUNDT Kharal'd (DE),
POPP Kharal'd (DE)**

(73) Proprietor(s):

**FRAUNKHOFER-GEZELL'SHAFT TsUR
FERDERUNG DER ANGEVANDTEN
FORSHUNG E.F. (DE)**

(54) DEVICE AND METHOD FOR GENERATING ENCODED STEREO SIGNAL OF AUDIO PART OR STREAM OF AUDIO DATA

(57) Abstract:

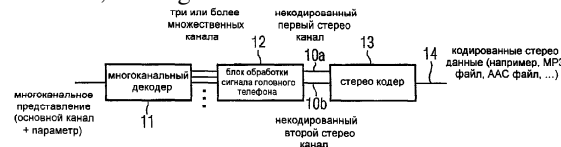
FIELD: physics; communications.

SUBSTANCE: invention relates to technology of multichannel audio and, specifically, to applications of multichannel audio in connections with headphone technologies. The device for generating an encoded stereo signal from a multichannel presentation includes a multichannel decoder (11), which forms three or more channels from at least one main channel and parametric information. Said three or more channels are subject to processing (12) headphone signals so as to generate an uncoded first stereo channel and an uncoded second stereo channel, which are then input into a stereo encoder (13) so as to generate an encoded stereo file at the

output side. The encoded stereo file can be transmitted to any suitable playback device in form of a CD player or portable playback device such that, the user not only receives a normal stereo impression, but a multichannel impression as well.

EFFECT: efficient signal processing concept, which allows for multichannel quality playback on headphones in simple playback devices.

12 cl, 11 dwg



Фиг. 1

Настоящее изобретение относится к технологии многоканального аудио и, в частности, к приложениям многоканального аудио в соединении с технологиями головных телефонов.

Международные публикации патентных заявок WO 99/49574 и WO 99/14983 раскрывают технологии обработки аудиосигнала для возбуждения пары противоположно размещаемых громкоговорителей головного телефона (наушников) для того, чтобы пользователь получил пространственное восприятие аудиосцены через эти два головных телефона, которое является не только стереопредставлением, но и многоканальным представлением. Таким образом, слушатель получит с помощью его или ее головных телефонов пространственное восприятие аудиочасти, которая в лучшем случае эквивалентна его или ее пространственному восприятию, если бы пользователь находился в зале воспроизведения, который, например, оборудован аудиосистемой 5.1. Для этой цели для каждого громкоговорителя головного телефона каждый канал части многоканального аудио или многоканального потока данных аудио, как иллюстрируется на Фиг.2, снабжен отдельным фильтром, вследствие чего соответствующие фильтрованные каналы, относящиеся друг к другу, суммируются, как будет проиллюстрировано ниже.

Слева на Фиг.2 имеются многоканальные входы 20, которые вместе представляют многоканальное представление аудиочасти или потока данных аудио. Такой сценарий, например, схематично показан на Фиг.10. Фиг.10 изображает пространство 200 воспроизведения, в котором размещается так называемая аудиосистема 5.1. Аудиосистема 5.1 включает в себя центральный громкоговоритель 201, передний левый громкоговоритель 202, передний правый громкоговоритель 203, задний левый громкоговоритель 204 и задний правый громкоговоритель 205. Аудиосистема 5.1 содержит дополнительный сабвуфер 206, который также называется как низкочастотный канал расширения. В так называемом "пятне благозвучного воспроизведения" пространства 200 воспроизведения находится слушатель 207, носящий головной телефон 208, содержащий левый громкоговоритель 209 головного телефона (наушника) и правый громкоговоритель 210 головного телефона (наушника).

Средство обработки, показанное на Фиг.2, сформировано так, чтобы фильтровать каждый канал 1, 2, 3 многоканальных входных сигналов 20 фильтром H_{iL} , описывающим звуковой канал от громкоговорителя на левый громкоговоритель 209 на Фиг.10, и дополнительно фильтровать тот же самый канал фильтром H_{iR} , представляющим звук от одного из этих пяти громкоговорителей к правому уху или правому громкоговорителю 210 головного телефона 208.

Если, например, канал 1 на Фиг.2 был передним левым каналом, выдаваемым громкоговорителем 202 на Фиг.10, фильтр H_{iL} будет представлять канал, обозначенный прерывистой линией 212, в то время как фильтр H_{iR} будет представлять канал, обозначенный прерывистой линией 213. Как в качестве примера обозначено на Фиг.10 прерывистой линией 214, левый громкоговоритель 209 головного телефона не только принимает непосредственный (излученный прямо) звук, но также и ранние отражения на границе пространства воспроизведения и, конечно, также поздние отражения, выраженные в диффузной реверберации.

Такое представление фильтра проиллюстрировано на Фиг.11. В частности, Фиг.11 изображает схематический пример импульсной характеристики фильтра, такого как, например, фильтр H_{iL} на Фиг.2. Непосредственный или первичный звук,

проиллюстрированный на Фиг.11 линией 212, представлен пиком в начале фильтра, в то время как ранние отражения, как проиллюстрировано в качестве примера на Фиг.10 позицией 214, воспроизводятся центральной областью, имеющей несколько (дискретных) маленьких пиков на Фиг.11. Диффузная реверберация обычно больше не разрешается для индивидуальных пиков, так как звук громкоговорителя 202 в принципе отражается произвольно часто, причем энергия конечно уменьшается с каждым отражением и дополнительным путем распространения, как проиллюстрировано уменьшающейся энергией в задержанной части, которая на Фиг.11 названа как "Диффузная реверберация".

Каждый фильтр, показанный на Фиг.2, таким образом включает в себя импульсную характеристику фильтра, грубо имеющую профиль (очертание), как показано схематической иллюстрацией импульсной характеристики на фиг.11. Очевидно, что импульсная характеристика отдельного фильтра будет зависеть от пространства воспроизведения, позиционирования громкоговорителей, возможных особенностей ослабления в пространстве воспроизведения, например, из-за присутствия нескольких людей или из-за мебели в пространстве воспроизведения, и в идеале - также от характеристик отдельных громкоговорителей 201-206.

Тот факт, что сигналы всех громкоговорителей налагаются друг на друга в ухе слушателя 207, иллюстрируется сумматорами 22 и 23 на Фиг.2. Таким образом, каждый канал фильтруется соответствующим фильтром для левого уха, чтобы затем просто быть суммированными сигналами, выданными фильтрами, которые предназначены для левого уха, чтобы получить выходной сигнал головного телефона для левого уха L. По аналогии суммирование сумматором 23 для правого уха или правого громкоговорителя 210 головного телефона на Фиг.10 выполняется для того, чтобы получить выходной сигнал головного телефона для правого уха, посредством наложения друг на друга всех сигналов громкоговорителя, фильтрованных соответствующим фильтром, для правого уха.

Вследствие того факта, что кроме непосредственного звука имеются также ранние отражения и, в частности, диффузная реверберация, которая имеет особенно высокую важность для пространственного восприятия, для того чтобы тон не звучал синтетическим или "неуклюжим", но давал слушателю впечатление, что он или она фактически находится в концертном зале с его акустическими характеристиками, все импульсные характеристики отдельных фильтров 21 будут иметь значительные длительности. Свертка каждого отдельного множественного канала многоканального представления, имеющего два фильтра, уже приводит к значительной вычислительной задаче. Так как два фильтра требуются для каждого отдельного множественного канала, а именно один для левого уха и другой для правого уха, когда канал сабвуфера также обрабатывается отдельно, для воспроизведения многоканального представления 5.1 посредством головного телефона требуется общее количество из 12 полностью отличных фильтров. Все фильтры имеют, как становится очевидным из Фиг.11, очень длинную импульсную характеристику, чтобы быть способными не только рассматривать непосредственный звук, но также и ранние отражения и диффузную реверберацию, которая реально только придает аудиочасти правильное воспроизведение звука и хорошее пространственное впечатление.

Чтобы реализовать эту известную концепцию на практике, кроме устройства 220 многоканального воспроизведения, как показано на Фиг.10, требуется очень сложный блок 222 обработки виртуального звука, который обеспечивает сигналы для этих двух

громкоговорителей 209 и 210, представленные линиями 224 и 226 на Фиг.10.

Системы головных телефонов для формирования многоканального звука головного телефона являются сложными, большими и дорогими, что является следствием высокой вычислительной мощности, высоких текущих требований для 5 высокой вычислительной требуемой мощности и высоких требований к рабочей памяти для оценок импульсной характеристики, которые должны быть выполнены, и большого количества или дорогих элементов для устройства воспроизведения, подсоединенных к нему. Применения этого вида таким образом привязаны к 10 звуковым платам домашнего персонального компьютера, или звуковым платам портативных компьютеров, или домашним стереосистемам.

В частности, многоканальный звук головного телефона остается недоступным для непрерывно увеличивающегося рынка мобильных проигрывателей, таких как, 15 например, мобильные проигрыватели CD или, в частности, аппаратные устройства воспроизведения, так как требования к вычислениям для фильтрации множественных каналов, например, с 12 различными фильтрами не могут быть реализованы в этом ценовом сегменте ни в отношении ресурсов процессора, ни в отношении текущих требований к обычным устройствам с батарейным питанием. Это относится к 20 ценовому сегменту в основании (нижнем конце) ценовой шкалы.

Однако этот сам ценовой сегмент экономически является очень интересным из-за большого количества компонентов.

Задача настоящего изобретения состоит в том, чтобы обеспечить эффективную концепцию обработки сигналов, допускающую многоканальное качественное 25 воспроизведение на головном телефоне на простых устройствах воспроизведения.

Эта задача решается устройством для формирования закодированного стереосигнала согласно пункту 1 формулы изобретения, или способом для формирования закодированного стереосигнала согласно пункту 11 формулы 30 изобретения, или компьютерной программой согласно пункту 12 формулы изобретения.

Настоящее изобретение основано на обнаружении того факта, что высококачественный и привлекательный многоканальный звук головного телефона (наушника) может быть сделан доступным всем доступным проигрывателям, таким 35 как, например, проигрыватели CD или аппаратные устройства воспроизведения, подвергая многоканальное представление аудиочасти или потока данных аудио, то есть, например, представления 5.1 аудиочасти, обработке сигналов головного телефона вне аппаратного устройства воспроизведения, то есть, например, в 40 компьютере поставщика, имеющем большую вычислительную мощность. Согласно изобретению результат обработки сигналов головного телефона, однако, не просто воспроизводится, но подается на обычный стереокодер аудио, который затем формирует закодированный стереосигнал из левого канала головного телефона и правого канала головного телефона.

Этот кодированный стереосигнал может быть затем, подобно любому другому закодированному стереосигналу, не содержащему многоканальное представление, 45 подан на аппаратное устройство воспроизведения или, например, мобильный проигрыватель CD в форме CD. Устройство воспроизведения или проигрывания затем 50 обеспечит пользователя многоканальным звуком головного телефона без каких-либо дополнительных ресурсов или средств, которые должны быть добавлены к устройствам, уже существующим. Что является изобретенным, результат обработки сигналов головного телефона, то есть сигнал левого и правого головного телефона

(наушника), не воспроизводится в головном телефоне, как это имеет место согласно уровню техники, а кодируется и выводится в качестве закодированных стереоданных.

Такие выходные данные могут быть сохранены, переданы и т.п. Такой файл, имеющий закодированные стереоданные, может затем легко быть подан на любое устройство воспроизведения, предназначенное для воспроизведения стерео, без необходимости для пользователя выполнять какие-либо изменения на своем устройстве.

Изобретательная концепция формирования закодированного стереосигнала на основе результата обработки сигналов головного телефона, таким образом, допускает многоканальное представление, обеспечивающее значительно улучшенное и более реальное качество для пользователя, которое также должно использоваться на всех простых и широко распространенных и в будущем даже более широко распространенных аппаратных устройствах воспроизведения.

В предпочтительном варианте осуществления настоящего изобретения отправной точкой является закодированное многоканальное представление, то есть параметрическое представление, содержащее один или обычно два основных канала и дополнительно содержащее параметрические данные, чтобы сформировать множественные каналы многоканального представления на основе основных каналов и упомянутых параметрических данных. Так как основанный на частотной области способ для многоканального декодирования является предпочтительным, обработка сигналов головного телефона согласно изобретению не выполняется во временной области посредством свертки временного сигнала с импульсной характеристикой, а выполняется в частотной области посредством умножения на функцию передачи фильтра.

Это допускает возможность сэкономить на, по меньшей мере, одном повторном преобразовании перед обработкой сигналов головного телефона и имеет то специфическое преимущество, когда последующий стереокодер также работает в частотной области, так что стереокодирование стереосигнала головного телефона без какой-либо необходимости перехода к временной области может также иметь место без перехода к временной области. Обработка из многоканального представления в закодированный стереосигнал без использования временной области или по меньшей мере уменьшенное количество преобразований является интересной не только в отношении эффективности времени вычислений, но и задает предел потерям качества, так как меньшее количество этапов обработки вводит меньшее количество артефактов в аудиосигнал.

В частности, в основанных на блочной обработке способах, выполняющих квантование, рассматривая психоакустический порог маскирования, что является предпочтительным для стереокодера, важно предотвратить как можно больше тандемных артефактов кодирования, насколько это возможно.

В особенно предпочтительном варианте осуществления настоящего изобретения представление ВСС, имеющее один или, предпочтительно, два основных канала, используется в качестве многоканального представления. Так как способ ВСС работает в частотной области, множественные каналы не преобразуются во временную область после синтеза, как обычно делается в декодере ВСС. Вместо этого спектральное представление множественных каналов в форме блоков используется и подвергается обработке сигналов головного телефона. Для этого используются функции преобразования фильтров, то есть преобразования Фурье импульсных характеристик, чтобы выполнить умножение спектрального представления

множественных каналов на функции преобразования фильтров. Когда импульсные характеристики фильтров во времени являются более длинными, чем блок спектральных компонентов на выходе декодера ВСС, предпочтительной является блочная обработка фильтра, когда импульсные характеристики фильтров отделяются
 5 во временной области и преобразуются блок за блоком, чтобы затем выполнить соответствующие взвешивания спектра, требуемые для метрик этого вида, как, например, раскрыто в публикации WO 94/01933.

Предпочтительные варианты осуществления настоящего изобретения подробно
 10 описаны ниже со ссылками на прилагаемые чертежи, на которых:

Фиг.1 показывает блок-схему предлагаемого устройства для формирования закодированного стереосигнала;

Фиг.2 является подробной иллюстрацией реализации обработки сигналов головного телефона согласно Фиг.1;

15 Фиг.3 показывает известный кодер объединенного стерео для формирования данных канала и параметрической многоканальной информации;

Фиг.4 является иллюстрацией схемы определения параметров ICLD, ICTD и ICC для кодирования/декодирования ВСС;

20 Фиг.5 является иллюстрацией блок-схемы цепочки кодер/декодер ВСС;

Фиг.6 изображает блок-схему реализации блока синтеза ВСС согласно Фиг.5;

Фиг.7 изображает каскадирование между многоканальным декодером и блоком обработки сигналов головного телефона без какого-либо преобразования к временной области;

25 Фиг.8 изображает каскадирование между блоком обработки сигналов головного телефона и стереокодером без какого-либо преобразования к временной области;

Фиг.9 изображает принципиальную блок-схему предпочтительного стереокодера;

Фиг.10 является принципиальной иллюстрацией сценария воспроизведения для
 30 определения функций фильтра согласно Фиг.2, и

Фиг.11 является принципиальной иллюстрацией ожидаемой импульсной характеристики фильтра, определенного согласно Фиг.10.

Фиг.1 показывает принципиальную блочную диаграмму предлагаемого устройства для формирования закодированного стереосигнала аудиочасти или потока данных
 35 аудио. Стереосигнал включает в себя, в некодированной форме, некодированный первый стереоканал 10a и некодированный второй стереоканал 10b и является сформированным из многоканального представления аудиочасти или потока данных аудио, при этом многоканальное представление содержит информацию относительно
 40 более чем двух множественных каналов. Как описано ниже, многоканальное представление может быть в некодированной или закодированной форме. Если многоканальное представление находится в некодированной форме, оно будет включать в себя три или более множественных каналов. В предпочтительном сценарии применения многоканальное представление включает в себя пять каналов и
 45 один канал сабвуфера.

Если многоканальное представление представлено, однако, в закодированной форме, эта закодированная форма будет обычно включать в себя один или несколько основных каналов, а также параметры для синтеза трех или более множественных
 50 каналов из одного или двух основных каналов. Многоканальный декодер 11, таким образом, является примером средства для обеспечения более чем двух множественных каналов из многоканального представления. Если многоканальное представление находится, однако, уже в некодированной форме, то есть, например, в форме 5+1

каналов ИКМ, это средство для обеспечения соответствует входному терминалу для средства 12 для выполнения обработки сигналов головного телефона, чтобы сформировать некодированный стереосигнал с некодированным первым стереоканалом 10a и некодированным вторым стереоканалом 10b.

Предпочтительно средство 12 для выполнения обработки сигналов головного телефона сформировано так, чтобы оценивать каждый множественный канал многоканального представления посредством первой функции фильтра для первого стереоканала и второй функции фильтра для второго стереоканала и суммировать соответствующие оцененные множественные каналы, чтобы получить некодированный первый стереоканал и некодированный второй канал стерео, как иллюстрируется со ссылками на Фиг.2. Следующим после средства 12 для выполнения обработки сигналов головного телефона является стереокодер 13, который сформирован так, чтобы кодировать первый некодированный стереоканал 10a и второй некодированный стереоканал 10b, чтобы получить закодированный стереосигнал на выходе 14 стереокодера 13. Стереокодер выполняет снижение скорости передачи данных, так что скорость передачи данных, требуемая для передачи закодированного стереосигнала, является меньшей, чем скорость передачи данных, требуемая для передачи некодированного стереосигнала.

Согласно изобретению достигается концепция, которая позволяет выдавать многоканальный тональный сигнал, который также назван как "окружающий" (surround), на стереоголовые телефоны (наушники) с помощью простых проигрывателей, таких как, например, аппаратные устройства воспроизведения.

Сумма некоторых каналов может в качестве примера быть сформирована как простая обработка сигналов головного телефона, чтобы получить выходные каналы для стереоданных. Усовершенствованные способы работают с более сложными алгоритмами, которые, в свою очередь, дают улучшенное качество воспроизведения.

Должно быть упомянуто, что изобретательная концепция позволяет не выполнять этапы интенсивных вычислений для многоканального декодирования и для выполнения обработки сигналов головного телефона непосредственно в устройстве воспроизведения, но выполнять внешним образом. Результатом изобретательной концепции является закодированный файл стерео, который является, например, MP3 файлом, AAC файлом, HE-AAC файлом или некоторым другим стереофайлом.

В других вариантах осуществления многоканальное декодирование, обработка сигналов головного телефона и кодирование стерео (сигнала) могут быть выполнены в различных устройствах, так как выходные данные и входные данные соответственно отдельных блоков могут быть легко перенесены и сформированы и сохранены стандартизированным способом.

Ниже ссылка делается на Фиг.7, иллюстрирующую предпочтительный вариант осуществления настоящего изобретения, где многоканальный декодер 11 содержит набор фильтров или функцию FFT (быстрого преобразования Фурье), так что обеспечивается многоканальное представление в частотной области. В частности, отдельные множественные каналы формируются как блоки спектральных значений для каждого канала. Изобретенным является то, что обработка сигналов головного телефона не выполняется во временной области посредством свертки временных каналов с импульсными характеристиками фильтра, но выполняется перемножение представления в частотной области множественных каналов на спектральное представление импульсной характеристики фильтра. Некодированный стереосигнал получают на выходе (блока) обработки сигналов головного телефона, который,

однако, представлен не во временной области, а включает в себя левый и правый каналы стерео, при этом такой стереоканал задается как последовательность блоков спектральных значений, причем каждый блок спектральных значений представляет кратковременный спектр стереоканала.

В варианте осуществления, показанном на Фиг.8, на вход блока 12 обработки сигналов головного телефона подаются данные или временной области, или частотной области. На стороне выхода некодированные каналы стерео формируются в частотной области, то есть снова в виде последовательности блоков спектральных значений. Стереокодер, который основан на преобразовании, то есть который обрабатывает спектральные значения без преобразования частота/время и последующего преобразования время/частота, необходим между блоком 12 обработки сигналов головного телефона и стереокодером 13, предпочтительно в качестве стереокодера 13 в этом случае. На стороне выхода стереокодер 13 затем выдает файл с закодированным сигналом стерео, который кроме дополнительной информации включает в себя закодированную форму спектральных значений.

В особенно предпочтительном варианте осуществления настоящего изобретения обработка в непрерывной частотной области выполняется на пути от многоканального представления на входе блока согласно Фиг.1 до закодированного стереофайла на выходе 14 из средства на Фиг.1 без преобразования во временную область и, возможно, обратного преобразования в частотную область, если в этом имеется необходимость. Когда кодер MP3 или кодер AAC используются в качестве стереокодера, будет предпочтительно преобразовать спектр Фурье на выходе блока обработки сигналов головного телефона к MDCT (модифицированное косинусное преобразование) спектру. Таким образом, согласно изобретению обеспечивается, что информация фазы, требуемая в точной форме для свертки/оценки каналов в блоке обработки сигналов головного телефона, преобразуется к MDCT представлению, не работающему в таком фазово-корректном способе, так что средство для преобразования из временной области в частотную область, то есть к MDCT спектру, не требуется для стереокодера в отличие от обычного MP3-кодера или обычного AAC-кодера.

Фиг.9 изображает общую блок-схему предпочтительного стереокодера. Стереокодер включает в себя, со стороны входа, модуль 15 объединенного стерео, который предпочтительно определяет адаптивным способом, обеспечивает ли обычное стереокодирование, например, в форме кодирования центрального/боковых каналов, более высокое усиление, чем отдельная обработка левых и правых каналов. Модуль 15 объединенного стерео может быть также сформирован так, чтобы выполнить стереокодирование интенсивности, в котором стереокодирование интенсивности, в частности, с верхними частотами, обеспечивает значительное усиление без возникновения слышимых артефактов. Выводной сигнал модуля 15 объединенного стерео затем обрабатывается, дополнительно используя различные другие сокращающие избыточность меры, такие как, например, TNS фильтрование, замещение шума и т.д., чтобы затем подать результаты на квантователь 16, который выполняет квантование спектральных значений, используя психоакустический порог маскирования. Размер шага квантователя здесь выбирают таким, что шум, введенный квантованием, остается ниже психоакустического порога маскирования, так что сокращение скорости передачи данных достигается без искажений, введенных квантованием с потерями, чтобы быть слышимыми. После квантователя 16 имеется статистический кодер 17, выполняющий статистическое кодирование без потерь

квантованных спектральных значений. На выходе статистического кодера имеется закодированный стереосигнал, который кроме статистически кодированных спектральных значений включает в себя дополнительную информацию, требуемую для декодирования.

Ниже ссылка делается на предпочтительные варианты реализации многоканального декодера и на предпочтительные иллюстрации множественных каналов, используя Фиг.3-6.

Имеется несколько способов для сокращения количества данных, требуемых для передачи многоканального аудиосигнала. Такие способы также называются способами объединенного стерео. Для этой цели ссылка делается на Фиг.3, иллюстрирующую устройство 60 объединенного стерео. Это устройство может быть устройством, реализующим, например, методику стереоинтенсивности (IS) или методику кодирования бинаурального (стерео) сигнала (BCC). Такое устройство обычно принимает по меньшей мере два канала CH_1 , CH_2 , ..., CH_n в качестве входного сигнала и выдает единственный канал-носитель и параметрическую многоканальную информацию. Параметрические данные определены так, чтобы аппроксимация первоначального канала (CH_1 , CH_2 , ..., CH_n) могла быть вычислена в декодере.

Обычно канал-носитель будет включать в себя выборки поддиапазона, спектральные коэффициенты, выборки во временной области и т.д., которые обеспечивают относительно точное представление основного сигнала, в то время как параметрические данные не включают в себя такие выборки или спектральные коэффициенты, но управляют параметрами для управления некоторым алгоритмом восстановления, таким как, например, взвешивание перемножением, смещение во времени, смещение по частоте и т.д. Параметрическая многоканальная информация таким образом включает в себя относительно грубое представление сигнала или связанного канала. Выражая в числах, количество данных, требуемых каналом-носителем, находится в диапазоне от 60 до 70 кбит/сек, в то время как количество данных, требуемых параметрической дополнительной информацией для канала, находится в диапазоне от 1,5 до 2,5 кбит/сек. Должно быть упомянуто, что вышеупомянутые числа относятся к сжатым данным. Несжатый канал CD конечно требует приблизительно десятикратных скоростей передачи данных. Примером параметрических данных являются известные масштабные коэффициенты, информация стереоинтенсивности или параметры BCC, как описано ниже.

Методика кодирования интенсивности стерео описана в AES Preprint 3799 "Intensity Stereo Coding" J. Herre, K.H. Brandenburg, D. Lederer, февраль 1994, Амстердам. Вообще концепция стереоинтенсивности основана на преобразовании основной оси, которая должна применяться к данным от двух стереофонических аудиоканалов. Если большинство точек данных сконцентрировано вокруг первой основной оси, усиление кодирования может быть достигнуто посредством вращения обоих сигналов на некоторый угол прежде, чем кодирование будет иметь место. Однако это не всегда относится к способам воспроизведения реального стерео. Таким образом, эта методика модифицируется в том, что второй ортогональный компонент исключается из передачи в потоке битов. Таким образом, восстановленные сигналы для левого и правого каналов состоят из различным образом взвешенных или масштабированных версий одного и того же переданного сигнала. Тем не менее, восстановленные сигналы отличаются по амплитуде, но они идентичны относительно их фазовой информации. Огибающие времени энергии обоих первоначальных аудиоканалов,

однако, поддерживается посредством избирательной операции масштабирования, обычно работающей частотно-избирательным способом. Это соответствует человеческому восприятию звука на высоких частотах, где доминирующая пространственная информация определяется огибающими энергии.

Кроме того, при практической реализации переданный сигнал, то есть канал-носитель, формируется из суммарного сигнала левого канала и правого канала вместо вращения обоих компонентов. Дополнительно эта обработка, то есть формирование параметров стереоинтенсивности для выполнения операций масштабирования, выполняется частотно-избирательным способом, то есть независимо для каждого диапазона масштабного коэффициента, то есть для каждого сегмента частоты кодера. Предпочтительно, оба канала объединяются, чтобы сформировать объединенный или канал "носитель" и, в дополнение к объединенному каналу, информацию стереоинтенсивности. Информация стереоинтенсивности зависит от энергии первого канала, энергии второго канала или энергии объединенного канала.

Методика BCC описана в AES Convention Paper 5574 "Binaural Cue Coding applied to stereo and multichannel audio compression" T. Faller, F. Baumgarte, May 2002, Мюнхен.

При BCC кодировании множество входных каналов аудио преобразовывают в спектральное представление, используя основанное на DFT (дискретное преобразование Фурье) преобразование с перекрывающимися окнами. Результирующий спектр разделяется на не перекрывающиеся части (сегменты), из которых каждая имеет индекс. Каждый сегмент имеет ширину полосы, которая является пропорциональной, эквивалентной прямоугольной ширине полосы частот (ERB). Межканальные разности уровней (ICLD) и межканальные разности во времени (ICTD) определяются для каждого сегмента и для каждого кадра k. ICLD и ICTD квантуются и кодируются, чтобы в конечном итоге получить поток битов BCC в качестве дополнительной информации. Межканальные разности уровней и межканальные разности во времени задаются для каждого канала по отношению к опорному (эталонному) каналу. Затем параметры вычисляют согласно заранее определенным формулам в зависимости от конкретных сегментов сигнала, который должен быть обработан.

На стороне декодера декодер обычно принимает моносигнал и поток битов BCC. Моносигнал преобразуется в частотную область и подается в блок пространственного синтеза, который также принимает декодированные значения ICLD и ICTD. В блоке пространственного синтеза параметры BCC (ICLD и ICTD) используются, чтобы выполнить операцию взвешивания моносигнала, синтезировать многоканальные сигналы, которые после преобразования частота/время представляют реконструкцию первоначального многоканального аудиосигнала.

В случае BCC модуль 60 объединенного стерео работает для вывода информации дополнительного канала, так что параметрические каналные данные являются квантованными и кодированными параметрами ICLD или ICTD, причем один из первоначальных каналов используется как опорный канал для кодирования информации дополнительного канала.

Обычно сигнал-носитель формируют из суммы участвующих первоначальных каналов.

Вышеупомянутые способы, конечно, только обеспечивают монофоническое представление для декодера, который может только обрабатывать канал-носитель, но который не способен обработать параметрические данные для формирования одного

или нескольких приближений более чем одного входного канала.

Методика ВСС также описана в публикациях патентных заявок США US 2003/0219130 A1, US 2003/0026441 A1 и US 2003/0035553 A1. Дополнительно ссылка может быть сделана на экспертную публикацию "Binaural Cue Coding. Part II Schemes and Applications" Т. Faller и F. Baumgarte, IEEE Trans. On Audio and Speech Proc., Vol. 11, № 6, ноябрь 2003.

Ниже обычная схема ВСС многоканального аудиокодирования проиллюстрирована более подробно со ссылками на Фиг.4-6.

Фиг.5 изображает такую схему ВСС для кодирования/передачи многоканальных аудиосигналов. Входной сигнал многоканального аудио на входе 110 кодера ВСС 112 смешивается с уменьшением числа каналов в так называемом блоке 114 смешения с уменьшением. В этом примере первоначальный многоканальный сигнал во входе 110 является 5-канальным сигналом окружающего (surround) звука, имеющим передний левый канал, передний правый канал, левый канал окружающего звука, правый канал окружающего звука и центральный канал. В предпочтительном варианте осуществления настоящего изобретения блок смешения с уменьшением формирует суммарный сигнал посредством простого добавления этих пяти каналов в один моносигнал.

Другие схемы смешения с уменьшением известны в области техники, так что, используя многоканальный входной сигнал, получают канал смешения с уменьшением, имеющий единственный канал.

Этот единственный канал выводится на линию 115 суммарного сигнала.

Дополнительная информация, полученная из блока анализа ВСС, выводится на линию 117 дополнительной информации.

Межканальные разности уровней (ICLD) и межканальные разности во времени (ICTD) вычисляются в блоке анализа ВСС, как было проиллюстрировано выше. Теперь блок 116 анализа ВСС также способен вычислить значения межканальной корреляции (значения ICC). Суммарный сигнал и дополнительная информация передаются к декодеру 120 ВСС в квантованном и закодированном формате. Декодер ВСС разделяет переданный суммарный сигнал на множество поддиапазонов и выполняет масштабирование, вводит задержки и дальнейшие этапы обработки, чтобы обеспечить поддиапазоны многоканальных аудиоканалов, которые нужно вывести. Эта обработка выполняется так, что параметры (сигналы) ICLD, ICTD и ICC восстановленного многоканального сигнала на выходе 121 совпадают с соответствующими сигналами для первоначального многоканального сигнала на входе 110 в кодер ВСС 112. Для этой цели декодер ВСС 120 включает в себя блок синтеза ВСС и блок обработки дополнительных данных.

Ниже внутренняя установка блока синтеза ВСС иллюстрируется со ссылками на Фиг.6. Суммарный сигнал на линии 115 подается в модуль преобразования время/частота или блок 125 фильтров FB (БФ). На выходе блока 125 имеется количество N сигналов поддиапазонов или, в крайнем случае, блок спектральных коэффициентов, когда блок 125 аудиофильтров выполняет преобразование 1:1, то есть преобразование, формирующее N спектральных коэффициентов из N выборок временной области.

Блок 122 синтеза ВСС также включает в себя каскад 126 задержки, каскад 127 модификации уровня, каскад 128 обработки корреляции 128 и каскад 129 инверсных фильтров IFB (КИФ). На выходе каскада 129 восстановленный многоканальный аудиосигнал, имеющий, например, пять каналов в случае 5-канальной системы

окружающего звука, может быть выведен на набор 124 громкоговорителей, как проиллюстрировано на Фиг.5 или Фиг.4.

Входной сигнал sn преобразуется в частотную область или область блока фильтров посредством элемента 125. Сигнал, выданный элементом 125, копируется так, что получают несколько версий одного и того же сигнала, как иллюстрируется узлом 130 копирования. Количество версий исходного сигнала равняется количеству выходных каналов в выходном сигнале. Затем каждая версия исходного сигнала в узле 130 подвергается некоторой задержке $d_1, d_2, \dots, d_1, \dots, d_N$. Параметры задержки вычисляются блоком 123 обработки дополнительных данных на Фиг.5 и получаются из межканальных разностей во времени, когда они были вычислены блоком 116 анализа ВСС согласно Фиг.5.

То же самое применяется к параметрам $a_1, a_2, \dots, a_1, \dots, a_N$ умножения, которые также вычисляются блоком 123 обработки дополнительной информации на основании межканальных разностей уровней, когда они были вычислены блоком 116 анализа ВСС.

Параметры ИСС, вычисленные блоком 116 анализа ВСС, используются для управления функциональными возможностями блока 128 так, чтобы некоторые корреляции между задержанными сигналами и сигналами с управляемым уровнем были получены на выводах блока 128. Должно быть отмечено, что последовательность этапов 126, 127, 128 может отличаться от последовательности, показанной на Фиг.6.

Должно быть также отмечено, что в основанной на кадре обработке аудиосигнала анализ ВСС также выполняют основанным на кадре, то есть изменяемым во времени, и что дополнительный основанный на частоте анализ ВСС получают, как может быть замечено, делением блока фильтров на Фиг.6. Это означает, что параметры ВСС получают для каждого спектрального диапазона. Это также означает, что в случае, когда блок 125 аудиофильтров разбивает входной сигнал на, например, 32 сигнала диапазонов, блок анализа ВСС получает набор параметров ВСС для каждого из этих 32 диапазонов. Конечно, блок 122 синтеза ВСС согласно Фиг.5, который проиллюстрирован более подробно на Фиг.6, также выполняет восстановление, которое также является основанным, в качестве примера, на упомянутых 32 диапазонах.

Ниже сценарий, используемый для определения индивидуальных параметров ВСС, проиллюстрирован со ссылками на Фиг.4. Обычно параметры ICLD, ICTD и ИСС могут быть определены между парами каналов. Однако предпочтительно, чтобы параметры ICLD и ICTD были определены между опорным каналом и каждым другим каналом. Это иллюстрируется на Фиг.4А.

Параметры ИСС могут быть определены различными способами. Обычно параметры ИСС могут быть определены в кодере между всеми возможными парами каналов, как иллюстрируется на Фиг.4В. Имелось предложение вычислять только параметры ИСС между двумя самыми сильными каналами в любой момент времени, как иллюстрируется на Фиг.4С, который показывает пример, в котором вычисляется в любой момент времени параметр ИСС между каналами 1 и 2 и в другой момент времени вычисляется параметр ИСС между каналами 1 и 5. Декодер затем синтезирует межканальную корреляцию между самыми сильными каналами в декодере и использует некоторые эвристические правила для вычисления и синтеза межканальной когерентности для оставшихся пар каналов.

Относительно вычисления, например, параметров a_i умножения на основании

переданных параметров ICLD ссылка делается к AES Convention Paper № 5574. Параметры ICLD представляют распределение энергии первоначального многоканального сигнала. Без потери общности, предпочтительно, как показывается на Фиг.4А, брать 4 параметра ICLD, представляющие разности энергии между
 5 соответствующими каналами и передним левым каналом. В блоке 122 обработки дополнительной информации параметры $a_1, \dots, a_N$ умножения получают из параметров ICLD так, чтобы полная энергия всех восстановленных выходных каналов была одной и той же (или пропорциональной энергии переданного
 10 суммарного сигнала).

В варианте осуществления, показанном на Фиг.7, обходятся без преобразования частота/время, полученного блоками 129 инверсных фильтров IFB на Фиг.6. Вместо этого используются спектральные представления отдельных каналов на входе этих
 15 блоков обратных фильтров и подаются на устройство обработки сигналов головного телефона согласно Фиг.7, чтобы выполнить оценку отдельных множественных каналов с соответствующими двумя фильтрами в расчете на множественный канал без дополнительного преобразования частота/время.

В отношении полной обработки, имеющей место в частотной области, должно
 20 быть отмечено, что в этом случае многоканальный декодер, то есть, например, блок 125 фильтров, согласно Фиг.6, и стереокодер должны иметь одну и ту же разрешающую способность время/частота. Дополнительно, предпочтительно использовать один и тот же блок фильтров, который имеет, в частности,
 25 преимущество в том, что только единственный блок фильтров требуется для полной обработки, как иллюстрируется на Фиг.1. В этом случае результатом является особенно эффективная обработка, так как преобразования в многоканальном декодере и кодере стерео не должны быть вычислены.

Входные данные и выходные данные, соответственно, в изобретательной
 30 концепции являются, таким образом, предпочтительно закодированными в частотной области посредством банка преобразования/фильтра и закодированными согласно психоакустическим рекомендациям, используя эффекты маскирования, при этом, в частности, в декодере должно иметься спектральное представление сигналов. Примерами этого являются MP3 файлы, AAC файлы или AC3 файлы. Однако входные
 35 данные и выходные данные, соответственно, могут также быть закодированы, формируя сумму и разность, как имеет место в так называемых матричных процессах. Примерами этого являются Dolby ProLogic, LogicV или Circle Surround. Данные, в частности, многоканального представления могут быть дополнительно закодированы
 40 посредством параметрических способов, как имеет место в окружающем звуке MP3, в котором этот способ основан на методике BCC.

В зависимости от обстоятельств изобретательный способ для формирования может быть осуществлен в виде или аппаратного, или программного обеспечения. Реализация может быть на цифровом носителе данных, в частности, на диске или CD,
 45 имеющем сигналы управления, которые могут быть считаны с помощью электроники, что может быть выполнено совместно с программируемой компьютерной системой так, что способ будет выполнен. В целом изобретение также заключается в компьютерном программном продукте, имеющем программный код, сохраненный на
 50 машиночитаемом носителе для выполнения изобретательного способа, когда компьютерный программный продукт выполняется на компьютере. С другой стороны, изобретение может также быть реализовано в виде компьютерной программы, имеющей программный код для выполнения способа, когда

компьютерная программа выполняется на компьютере.

Формула изобретения

1. Устройство для формирования кодированного стереосигнала аудиочасти или потока данных аудио, имеющего первый стереоканал и второй стереоканал из многоканального представления аудиочасти или потока данных аудио, содержащего информацию относительно более двух множественных каналов, содержащее:

средство (11) для формирования, по меньшей мере, трех множественных каналов на основе одного или нескольких основных каналов и параметрической информации из многоканального представления аудиочасти или потока данных аудио, содержащего один или несколько основных каналов, а также параметрическую информацию;

средство (12) для выполнения обработки сигналов головного телефона для формирования некодированного стереосигнала с некодированным первым стереоканалом (10a) и некодированным вторым стереоканалом (10b), причем средство (12) для выполнения сформировано для того, чтобы

оценивать каждый множественный канал посредством первой функции фильтра (H_{iL}), полученной из виртуальной позиции громкоговорителя для воспроизведения этого множественного канала и виртуальной позиции первого уха слушателя, для первого канала стерео и второй функции фильтра (H_{iR}), полученной из виртуальной позиции громкоговорителя и виртуальной позиции второго уха слушателя, для второго канала стерео, чтобы сформировать первый оцененный канал и второй оцененный канал для каждого множественного канала, причем две виртуальные позиции уха слушателя являются различными,

суммировать (22) оцененные первые каналы, чтобы получить некодированный первый стереоканал (10a), и

суммировать (23) оцененные вторые каналы, чтобы получить некодированный второй стереоканал (10b); и

стерео кодер (13) для кодирования некодированного первого стереоканала (10a) и некодированного второго стереоканала (10b), чтобы получить закодированный стереосигнал (14), причем стереокодер сформирован так, что скорость передачи данных, требуемая для передачи закодированного стереосигнала, является меньшей, чем скорость передачи данных, требуемая для передачи некодированного стереосигнала.

2. Устройство по п.1, в котором средство (12) для выполнения сформировано для того, чтобы использовать первую функцию фильтра (H_{iL}), рассматривающую непосредственный звук, отражения и диффузную реверберацию, и вторую функцию фильтра (H_{iR}), рассматривающую непосредственный звук, отражения и диффузную реверберацию.

3. Устройство по п.2, в котором первая и вторая функции фильтра соответствуют импульсной характеристике фильтра, содержащей пик при малом значении времени, представляющий непосредственный звук, несколько меньших пиков при средних значениях времени, представляющих отражения, и непрерывную область, больше не разрешенную для отдельных пиков и представляющую диффузную реверберацию.

4. Устройство по п.1, в котором средство (11) для формирования сконфигурировано, чтобы сформировать на стороне выхода блочное представление в частотной области для каждого множественного канала, и

в котором средство (12) для выполнения сформировано для оценки блочного

представления в частотной области посредством представления в частотной области первой и второй функций фильтра.

5 5. Устройство по п.1, в котором средство (12) для выполнения сконфигурировано, чтобы сформировать блочное представление в частотной области некодированного первого стереоканала и некодированного второго канала стерео, и

в котором стерео кодер (13) является основанным на преобразовании кодером и также сформирован, чтобы обработать блочное представление в частотной области некодированного первого стереоканала и некодированного второго стереоканала без преобразования из представления в частотной области во временное представление.

10 6. Устройство по п.1, в котором стереокодер (13) сформирован так, чтобы выполнять обычное стерео кодирование (15) первого и второго каналов стерео.

7. Устройство по п.1, в котором стереокодер (13) сформирован так, чтобы квантовать (16) блок спектральных значений, используя психоакустический порог маскирования, и подвергать его статистическому кодированию (17), чтобы получить закодированный стереосигнал.

8. Устройство по п.1, в котором средство (11) для формирования сконфигурировано как декодер ВСС.

20 9. Устройство по п.1, в котором средство (11) для формирования сконфигурировано как многоканальный декодер, содержащий блок фильтров, имеющий несколько выходов,

при этом средство (12) для выполнения сформировано так, чтобы оценить сигналы на выходах блока фильтров первой и второй функциями фильтра, и

25 при этом стереокодер (13) сформирован так, чтобы квантовать (16) некодированный первый стереоканал в частотной области и некодированный второй канал стерео в частотной области и подвергать его статистическому кодированию (17), чтобы получить закодированный стереосигнал.

30 10. Способ формирования закодированного стереосигнала аудиочасти или потока данных аудио, имеющего первый стереоканал и второй стереоканал, из многоканального представления аудиочасти или потока данных аудио, содержащего информацию относительно более двух множественных каналов, содержащий этапы:

формируют (11), по меньшей мере, три множественных канала на основе одного или нескольких основных каналов и параметрической информации из многоканального представления аудиочасти или потока данных аудио, содержащего один или несколько основных каналов, а также параметрическую информацию;

35 выполняют (12) обработку сигналов головного телефона, чтобы сформировать некодированный стереосигнал с некодированным первым стереоканалом (10a) и некодированным вторым стереоканалом (10b), при этом этап выполнения (12) содержит этапы:

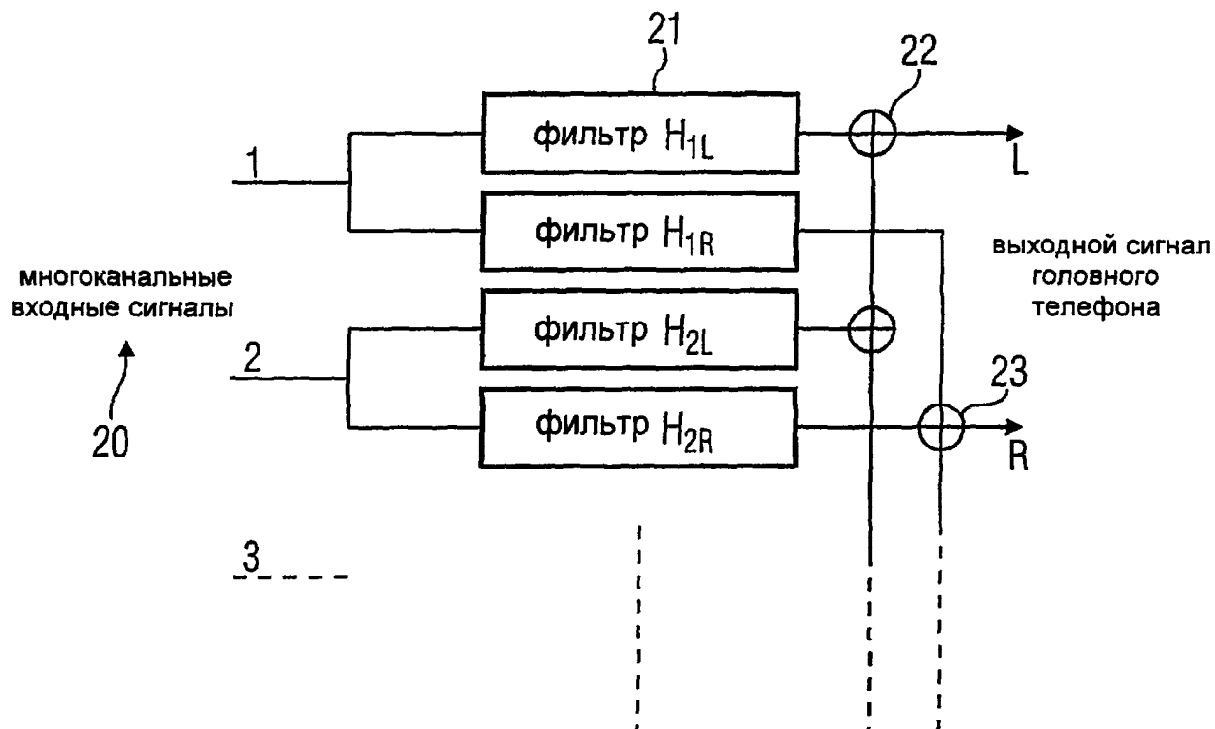
оценивают каждый множественный канал первой функцией фильтра (H_{iL}), полученной из виртуальной позиции громкоговорителя для воспроизведения 45 множественного канала и виртуальной позиции первого уха слушателя, для первого канала стерео и второй функцией фильтра (H_{iR}), полученной из виртуальной позиции громкоговорителя и виртуальной позиции второго уха слушателя, для второго канала стерео, чтобы сформировать первый оцененный канал и второй оцененный канал для каждого множественного канала, причем две виртуальные позиции уха слушателя 50 являются отличающимися,

суммируют (22) оцененные первые каналы для получения некодированного первого стереоканала (10a), и

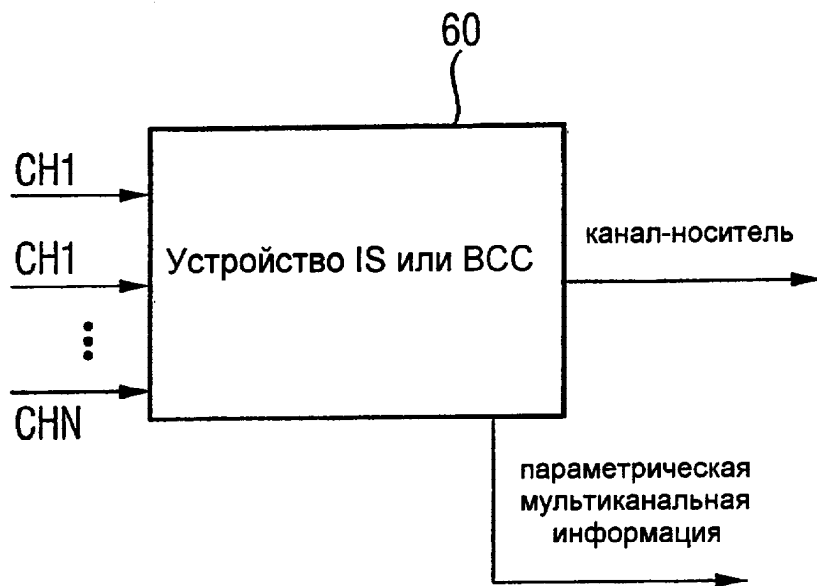
суммируют (23) оцененные вторые каналы, чтобы получить некодированный второй стереоканал (10b); и

выполняют стерео кодирование (13) некодированного первого стереоканала (10a) и некодированного второго стереоканала (10b), чтобы получить закодированный стереосигнал (14), при этом этап стереокодирования выполняется так, что скорость передачи данных, требуемая для передачи закодированного стереосигнала, является меньшей, чем скорость передачи данных, требуемая для передачи некодированного стереосигнала.

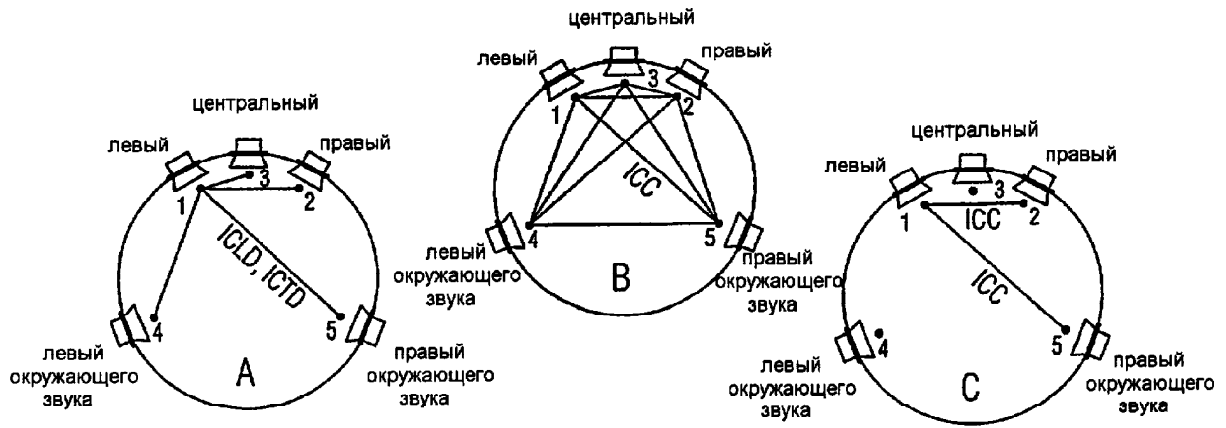
11. Машиночитаемый носитель информации, содержащий программируемый код, сохраненный на машиночитаемом носителе и исполняемый компьютером, и предназначенный для выполнения действий согласно способу по п.10.



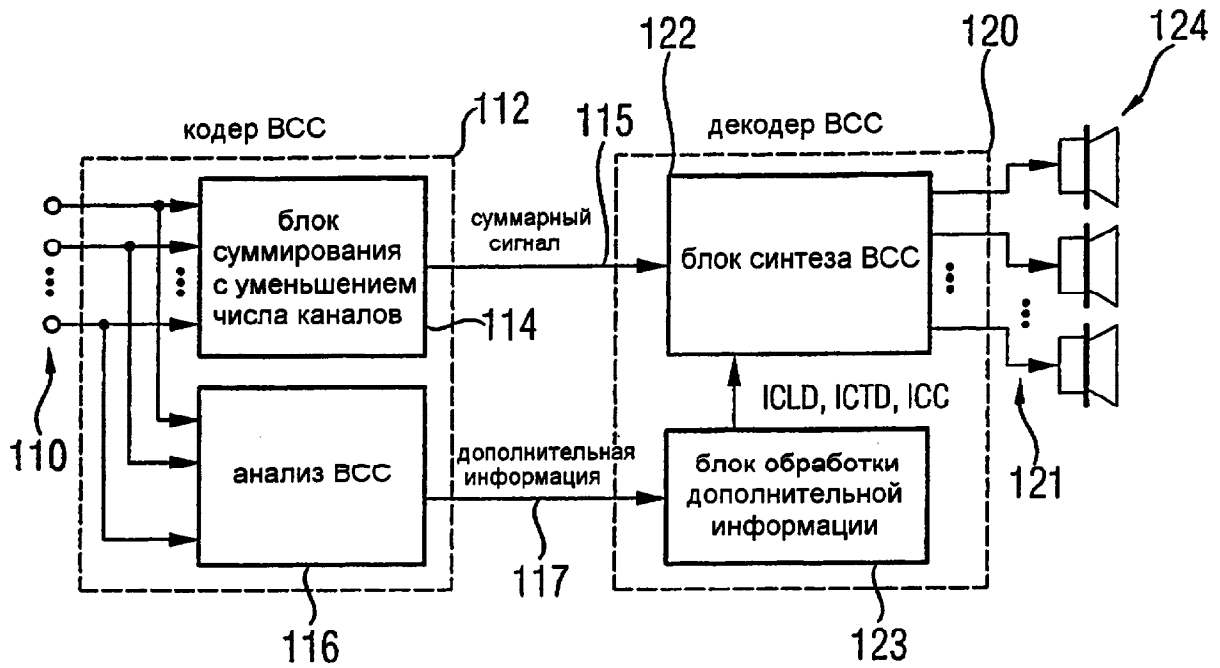
Фиг. 2



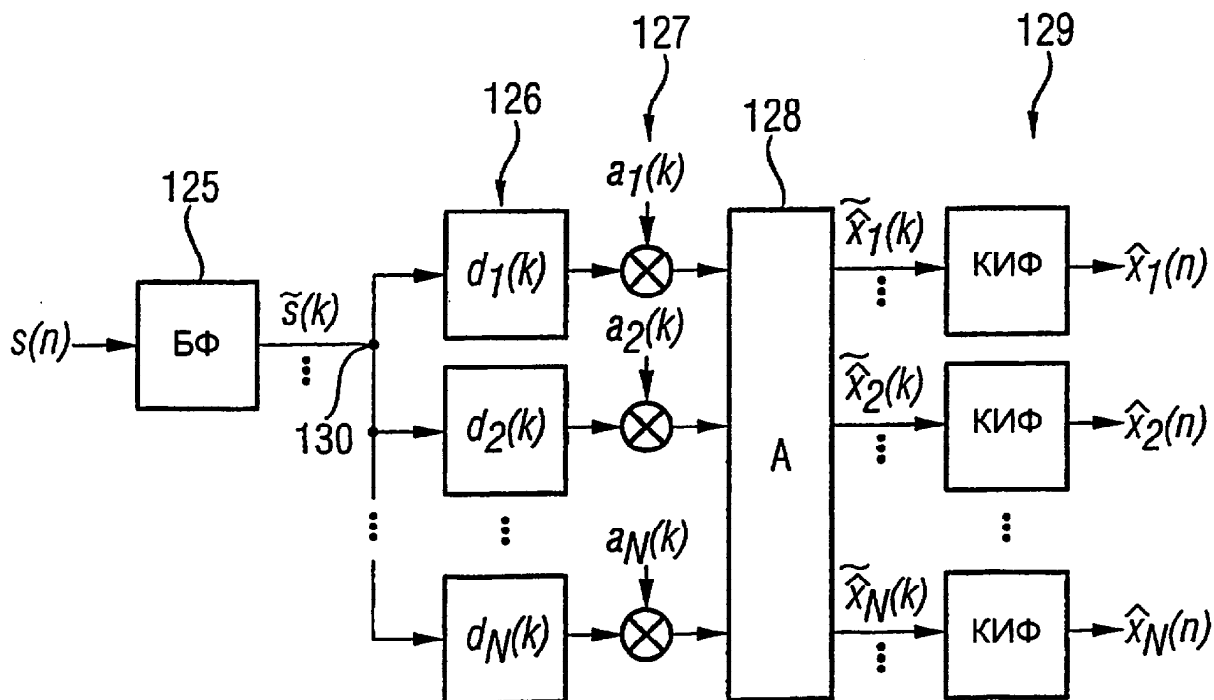
Фиг. 3



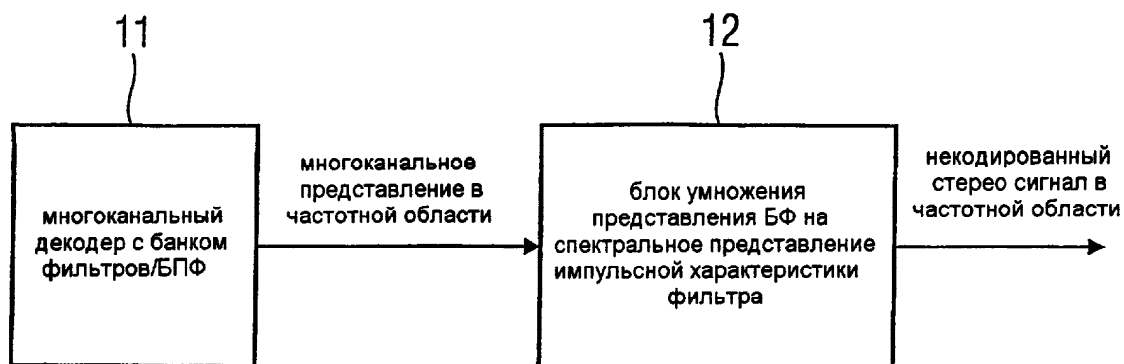
Фиг. 4



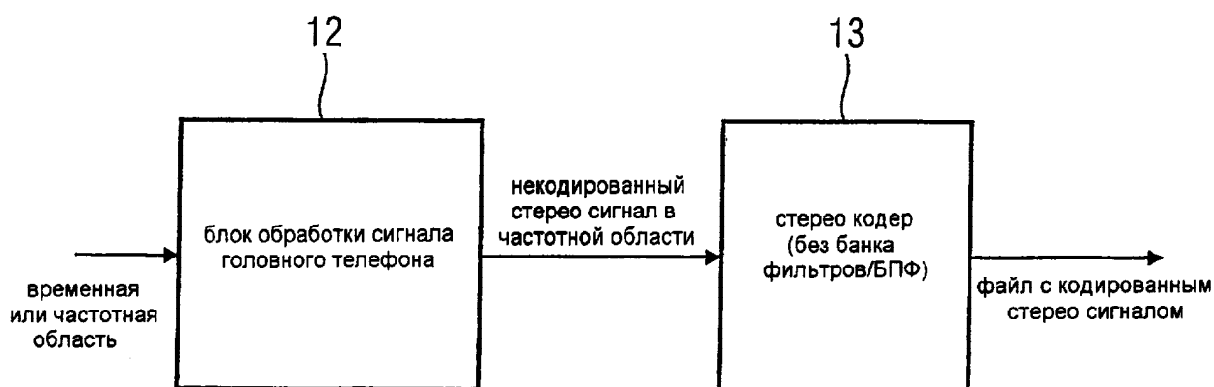
Фиг. 5



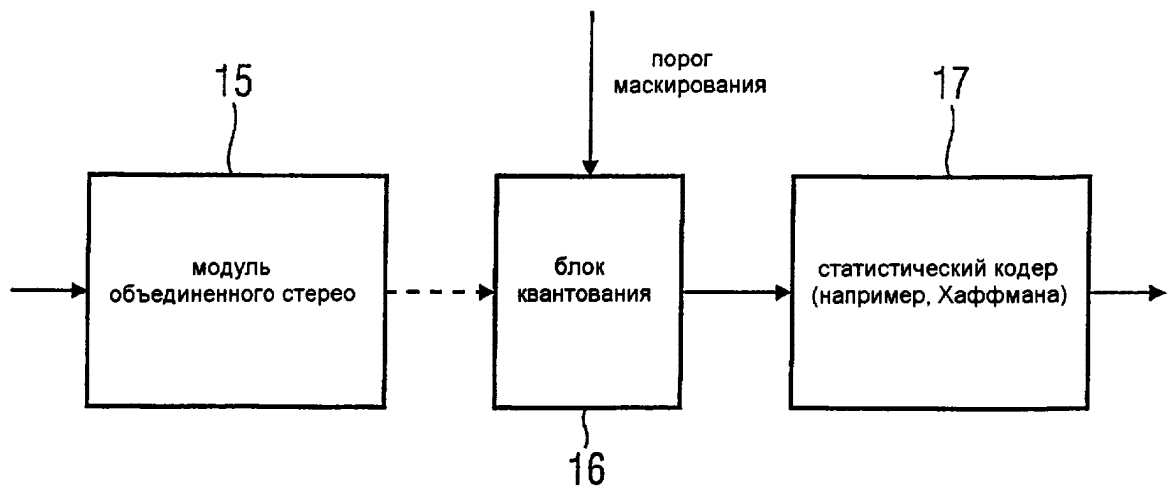
Фиг. 6



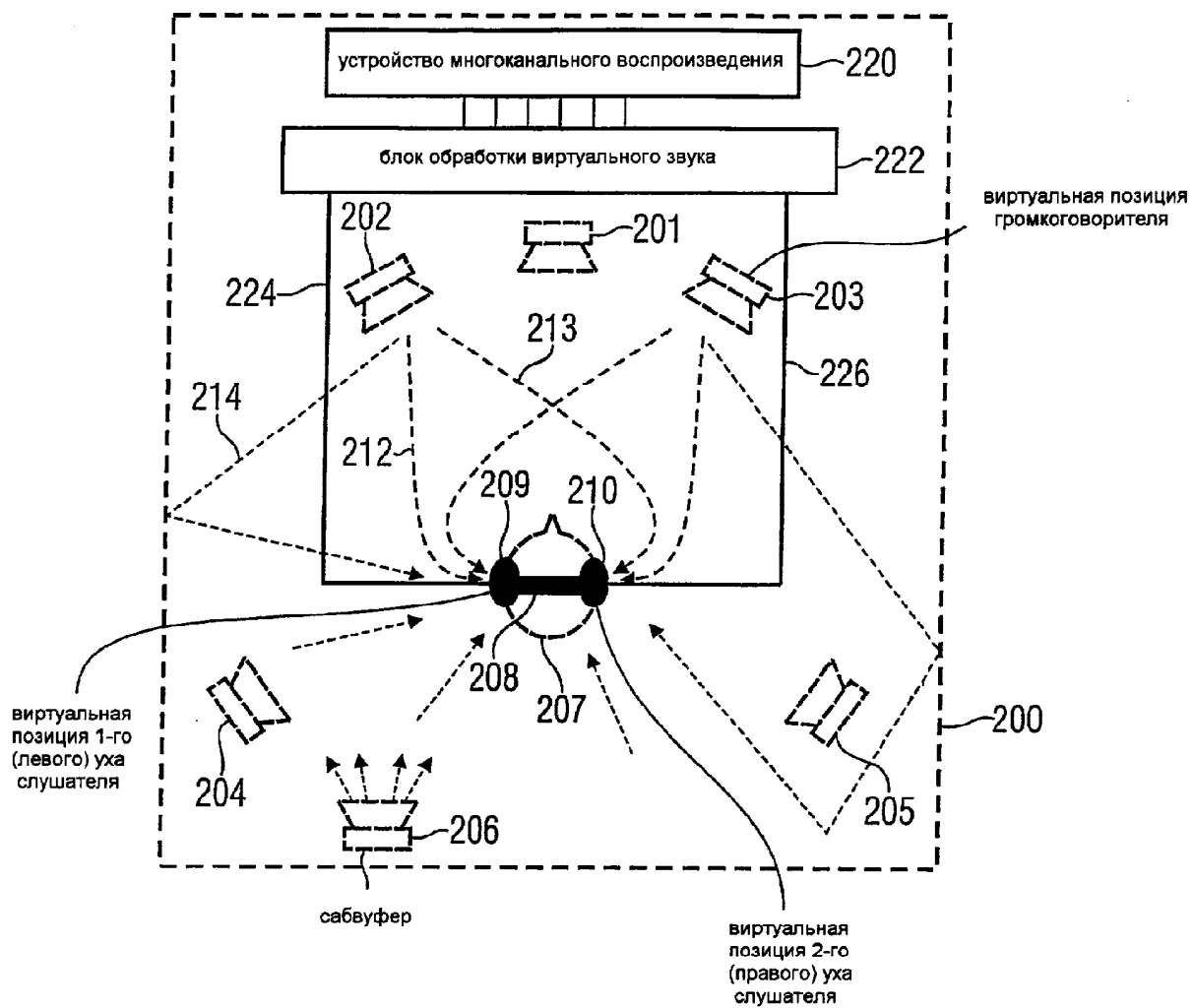
Фиг. 7



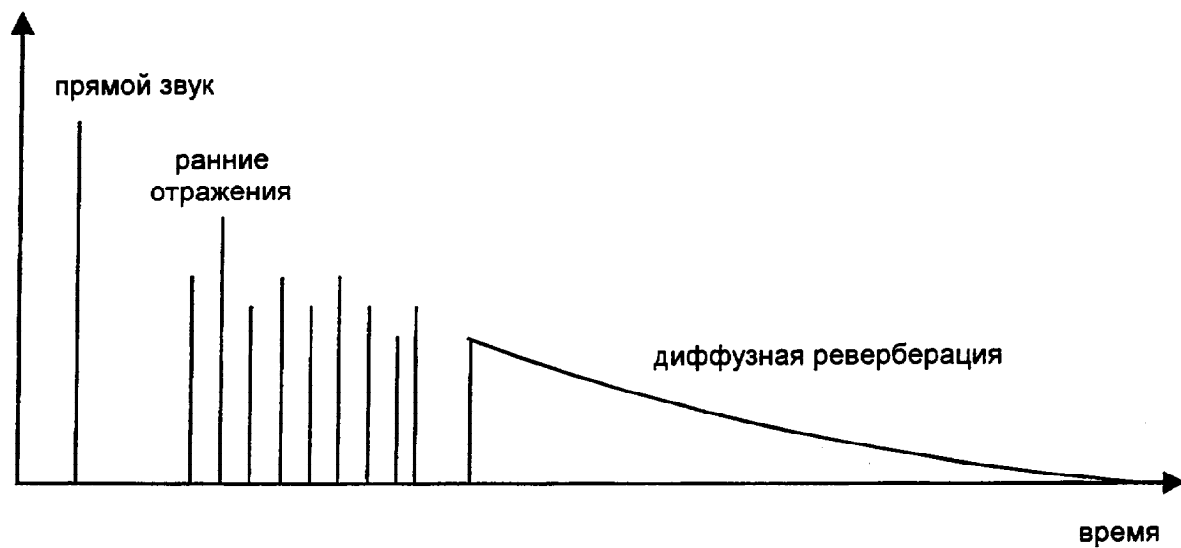
Фиг. 8



Фиг. 9



Фиг. 10



Фиг. 11