



ФЕДЕРАЛЬНАЯ СЛУЖБА
ПО ИНТЕЛЛЕКТУАЛЬНОЙ СОБСТВЕННОСТИ

(12) ОПИСАНИЕ ИЗОБРЕТЕНИЯ К ПАТЕНТУ

(52) СПК

H04S 7/30 (2024.08); H04S 7/302 (2024.08); H04S 7/303 (2024.08); H04S 7/307 (2024.08); H04S 2400/11 (2024.08); H04S 2420/01 (2024.08)

(21)(22) Заявка: 2024115701, 07.11.2022

(24) Дата начала отсчета срока действия патента:
07.11.2022Дата регистрации:
18.06.2025

Приоритет(ы):

(30) Конвенционный приоритет:
09.11.2021 EP 21207288.8

(43) Дата публикации заявки: 09.07.2024 Бюл. № 19

(45) Опубликовано: 18.06.2025 Бюл. № 17

(85) Дата начала рассмотрения заявки РСТ на
национальной фазе: 10.06.2024(86) Заявка РСТ:
EP 2022/080996 (07.11.2022)(87) Публикация заявки РСТ:
WO 2023/083752 (19.05.2023)

Адрес для переписки:

129090, Москва, ул. Большая Спасская, д.25,
стр.3, ООО "Юридическая фирма Городисский
и Партнеры"

(72) Автор(ы):

У, Юнь-Хань (DE),
ХЕРРЕ, Юрген (DE),
КОРОТАЕВ, Михаил (DE),
ГАЙЕР, Матиас (DE),
ШВЕР, Зимон (DE),
АДАМИ, Александер (DE),
АНЕМЮЛЛЕР, Карлотта (DE)

(73) Патентообладатель(и):

ФРАУНХОФЕР-ГЕЗЕЛЛЬШАФТ ЦУР
ФЕРДЕРУНГ ДЕР АНГЕВАНДТЕН
ФОРШУНГ Е.Ф. (DE)(56) Список документов, цитированных в отчете
о поиске: WO 2021180935 A1, 16.09.2021. US
2021084429 A1, 18.03.2021. WO 2013155217 A1,
17.10.2013. WO 2021144308 A1, 22.07.2021. WO
2020127329 A1, 25.06.2020. RU 2523215 C2,
20.07.2014. RU 2671627 C2, 02.11.2018.(54) УСТРОЙСТВО, СПОСОБ И КОМПЬЮТЕРНАЯ ПРОГРАММА ДЛЯ СИНТЕЗИРОВАНИЯ
ПРОСТРАНСТВЕННО ПРОТЯЖЕННОГО ИСТОЧНИКА ЗВУКА С ИСПОЛЬЗОВАНИЕМ
ЭЛЕМЕНТАРНЫХ ПРОСТРАНСТВЕННЫХ СЕКТОРОВ

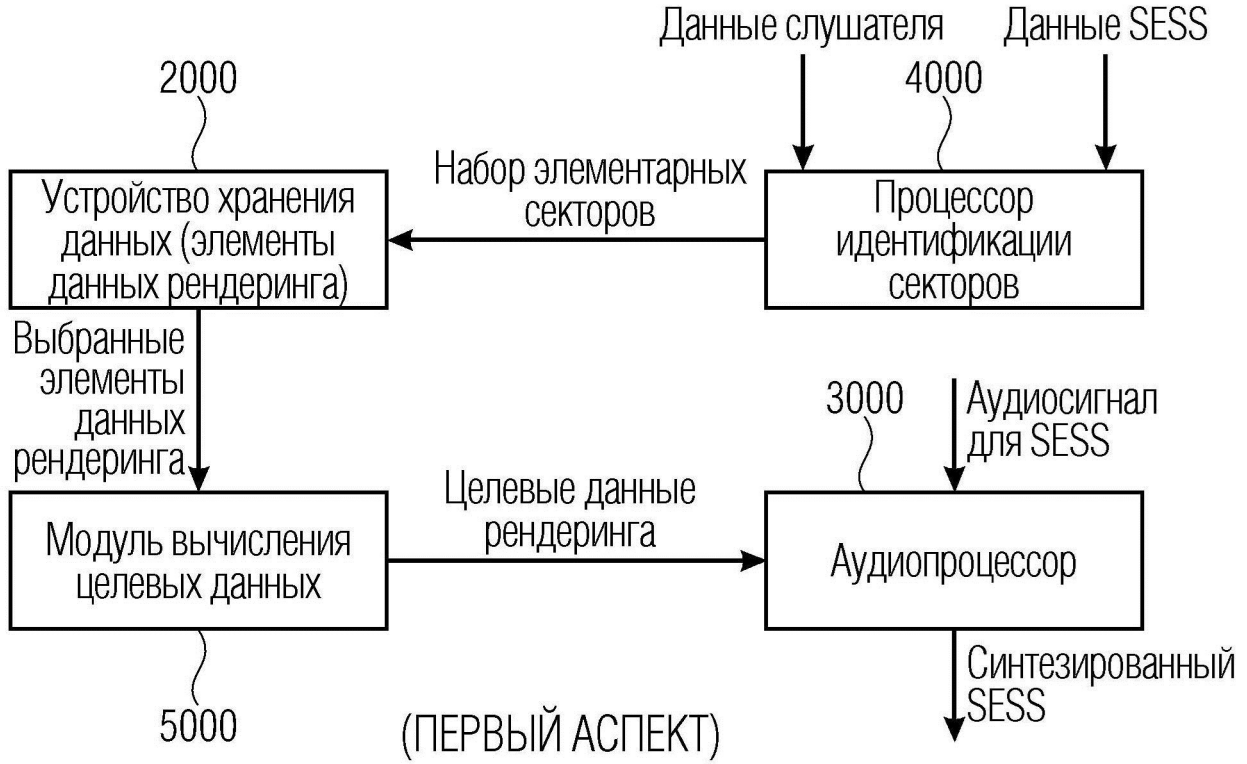
(57) Реферат:

Изобретение относится к обработке аудиосигналов и, в частности, относится к синтезу пространственно протяженных источников звука (SESS). Технический результат заключается в обеспечении возможности усовершенствованной концепции воспроизведения пространственно протяженного источника звука и возможности для пространственно избирательной модификации рендеринга SESS, в том числе с комплексной

геометрической формой. Устройство для синтеза пространственно протяженного источника (7000) звука (SESS) содержит устройство (200, 2000) хранения данных для сохранения элементов данных рендеринга для различных элементарных пространственных секторов, покрывающих диапазон рендеринга для слушателя, процессор (4000) идентификации секторов для идентификации набора

элементарных пространственных секторов, принадлежащих пространственно протяженному источнику звука, на основе данных слушателя и данных пространственно протяженных источников звука, модуль (5000) вычисления целевых данных для вычисления целевых данных рендеринга из элементов данных рендеринга для набора элементарных пространственных

секторов и аудиопроцессор (300, 3000) для обработки аудиосигнала, представляющего пространственно протяженный источник звука, с использованием целевых данных рендеринга. Заявлен также способ синтеза SESS, который реализуется в заявленном устройстве. 2 н. и 13 з.п. ф-лы, 14 ил.



ФИГ. 1

RU 2842007 C2

RU 2842007 C2



FEDERAL SERVICE
FOR INTELLECTUAL PROPERTY

(12) **ABSTRACT OF INVENTION**

(52) CPC

H04S 7/30 (2024.08); *H04S 7/302* (2024.08); *H04S 7/303* (2024.08); *H04S 7/307* (2024.08); *H04S 2400/11* (2024.08); *H04S 2420/01* (2024.08)

(21)(22) Application: **2024115701, 07.11.2022**

(24) Effective date for property rights:
07.11.2022

Registration date:
18.06.2025

Priority:

(30) Convention priority:
09.11.2021 EP 21207288.8

(43) Application published: **09.07.2024** Bull. № 19

(45) Date of publication: **18.06.2025** Bull. № 17

(85) Commencement of national phase: **10.06.2024**

(86) PCT application:
EP 2022/080996 (07.11.2022)

(87) PCT publication:
WO 2023/083752 (19.05.2023)

Mail address:
**129090, Moskva, ul. Bolshaya Spasskaya, d.25, str.3,
OOO "Yuridicheskaya firma Gorodisskij i
Partnery"**

(72) Inventor(s):

**WU, Yun-Han (DE),
HERRE, Juergen (DE),
KOROTIAEV, Mikhail (DE),
GEIER, Matthias (DE),
SCHWAER, Simon (DE),
ADAMI, Alexander (DE),
ANEMULLER, Carlotta (DE)**

(73) Proprietor(s):

**FRAUNHOFER-GESELLSCHAFT ZUR
FOERDERUNG DER ANGEWANDTEN
FORSCHUNG E.V. (DE)**

(54) **DEVICE, METHOD AND COMPUTER PROGRAM FOR SYNTHESIZING SPATIALLY EXTENDED SOUND SOURCE USING ELEMENTARY SPATIAL SECTORS**

(57) Abstract:

FIELD: audio signal processing.

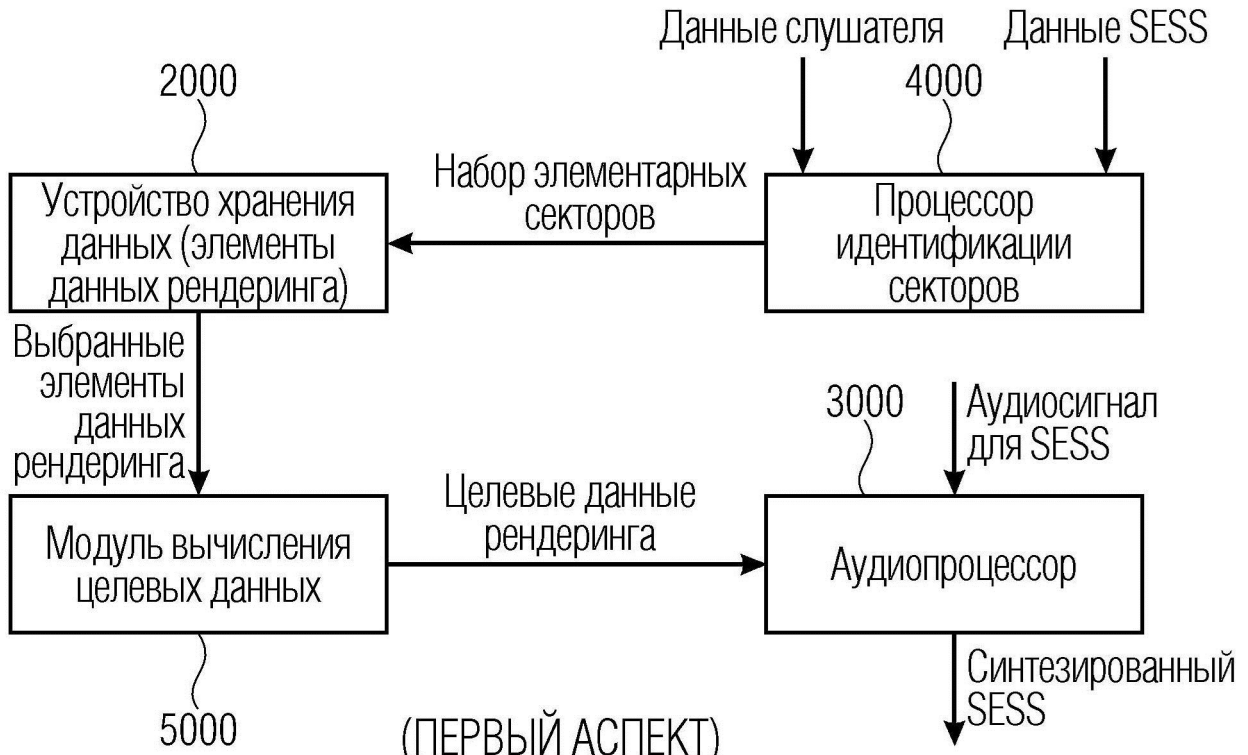
SUBSTANCE: invention relates to processing audio signals and particularly relates to synthesis of spatially extended sound sources (SESS). Device for synthesising spatially extended sound source (7000) (SESS) comprises data storage device (200, 2000) for storing rendering data elements for different elementary spatial sectors covering the rendering range for a listener, sector identification processor (4000) for identifying a set of elementary spatial sectors belonging to a spatially extended sound source based on the listener data and

the data of the spatially extended sound sources, a target data computing unit (5000) for calculating target rendering data from rendering data elements for a set of elementary spatial sectors and audio processor (300, 3000) for processing an audio signal representing a spatially extended sound source using the target rendering data. Also disclosed is a method for synthesis of SESS, which is implemented in disclosed device.

EFFECT: enabling an improved concept of reproducing a spatially extended sound source and enabling spatially selective modification of SESS

RU 2 842 007 C2

RU 2 842 007 C2



(ПЕРВЫЙ АСПЕКТ)
ФИГ. 1

RU 2842007 C2

RU 2842007 C2

Изобретение относится к обработке аудиосигналов и, в частности, относится к синтезу пространственно протяженных источников звука (SESS).

В течение длительного времени исследуется воспроизведение источников звука по нескольким громкоговорителям или наушникам. Простейший способ воспроизведения источников звука в таких компоновках состоит в том, чтобы рендерировать их в качестве точечных источников, т.е. очень (в идеале: бесконечно) небольших источников звука. Тем не менее, эта теоретическая концепция практически не позволяет моделировать существующие физические источники звука реалистичным способом. Например, рояль имеет большую вибрирующую деревянную крышку с множеством пространственно распределенных струн внутри и в силу этого кажется гораздо большим при слуховом восприятии, чем точечный источник (в частности, когда слушатель (и микрофоны) находятся близко к роялю). Множество источников звука реального мира имеют значительный размер («пространственную протяженность»), например, музыкальные инструменты, машины, оркестр или хор либо окружающие звуки (звук водопада).

Корректное/реалистичное воспроизведение таких источников звука становится задачей множества способов воспроизведения звука, независимо от того, являются ли они бинауральными (т.е. с использованием так называемых передаточных функций восприятия звука человеком (HRTF) или бинауральных импульсных откликов в помещении (BRIR)) с использованием наушников, либо традиционными с использованием компоновок громкоговорителей в пределах от 2 динамиков («стерео») до множества динамиков, расположенных в горизонтальной плоскости («объемный звук»), и множества динамиков, окружающих слушателя во всех трех измерениях («трехмерное аудио»).

В качестве примера, если SESS (например, фонтан) прослушивается из места, в котором часть фонтана загораживается посредством кустарников, загороженные части фонтана подвергаются процессу частотного демпфирования, т.е. ослабляются посредством определенного частотного отклика, который определяется посредством характеристик пропускания кустарника. Характеристики рендеринга таких (частично) загороженных частей SESS недоступны в первоначально описанном алгоритме рендеринга SESS. Аналогично, более удаленные части SESS могут реалистично рендерироваться с более низким уровнем с использованием настоящего изобретения.

Ширина двумерного источника

В этом разделе описаны способы, которые относятся к рендерингу протяженных источников звука на двумерной поверхности, обращенной с точки обзора слушателя, например, в определенном диапазоне изменения азимута при подъеме в нуль градусов (что имеет место в традиционном стерео-/объемном звуке), или в определенных диапазонах изменения азимута и подъема (что имеет место в трехмерном аудио или виртуальной реальности с 3 степенями свободы ["3DoF"]) перемещения пользователя, т.е. вращения головы в осях наклона в продольном направлении/наклона относительно вертикальной/крена).

Увеличение кажущейся ширины аудиообъекта, который панорамируется между двумя или более громкоговорителей (формирование так называемого фантомного изображения или фантомного источника) может достигаться посредством снижения корреляции участвующих канальных сигналов. Со снижением корреляции, разброс фантомных источников увеличивается до тех пор, пока, для корреляционных значений, близких к нулю (и не слишком широких углах раскрытия), он не охватывает весь диапазон между громкоговорителями.

Декоррелированные версии сигнала источника получаются посредством извлечения и применения подходящих декорреляционных фильтров. В работе автора Lauridsen предложено суммировать/вычитать масштабированную версию с временной задержкой сигнала источника относительно себя, чтобы получать две декоррелированных версии сигнала. Более сложные подходы, например, предложены в работе автора Kendall. Он итеративно извлекает спаренные декорреляционные всечастотные фильтры на основе комбинаций последовательностей случайных чисел. Работа авторов Faller и др. предлагает подходящие декорреляционные фильтры («рассеиватели»). Также в работе Zotter и др. извлекаются пары фильтров, в которых частотно-зависимые разности фаз или амплитуд использованы для достижения расширения фантомного источника. Кроме того, предложены декорреляционные фильтры на основе бархатного шума, которые дополнительно оптимизированы.

Помимо уменьшения корреляции соответствующих канальных сигналов фантомного источника, ширина источника также может увеличиваться посредством увеличения числа фантомных источников, приписанных аудиообъекту. В, ширина источника управляется посредством панорамирования одного и того же сигнала источника в (немного) различных направлениях. Первоначально предложен способ стабилизации воспринимаемого разброса фантомных источников сигналов панорамированных источников VBAP, когда они перемещаются в звуковой сцене. Это является преимущественным, поскольку в зависимости от направления источника, рендеримый источник воспроизводится посредством двух или более динамиков, что может приводить к нежелательным изменениям воспринимаемой ширины источника.

DirAC в виртуальном мире представляет собой расширение традиционного подхода на основе направленного кодирования аудио (DirAC) для синтеза звука в виртуальных мирах. Для рендеринга пространственной протяженности, направленные звуковые компоненты источника случайно панорамируются в пределах определенного диапазона вокруг исходного направления источника, причем направления панорамирования варьируются во времени и по частоте.

Аналогичный подход задействуется в, при котором пространственная протяженность достигается посредством случайного распределения полос частот сигнала источника для различных пространственных направлений. Он представляет собой способ, направленный на формирование пространственно распределенного и огибающего звука, поступающего одинаково из всех направлений, вместо управления точной степенью протяженности.

В работе авторов Verron и др. пространственная протяженность источника достигается не посредством использования панорамированных коррелированных сигналов, а посредством синтеза нескольких некогерентных версий сигнала источника, их равномерного распределения по окружности вокруг слушателя и смещения между собой. Число и усиление одновременно активных источников определяют интенсивность эффекта расширения. Этот способ реализован как пространственное протягивание для синтезатора для звуков окружающей среды.

Ширина трехмерного источника

В данном разделе описаны способы, которые относятся к рендерингу протяженных источников звука в трехмерном пространстве, т.е. объемным способом, что требуется для виртуальной реальности с 6 степенями свободы ("6DoF"). Это означает 6 степеней свободы перемещения пользователя, т.е. вращения головы в осях наклона в продольном направлении/наклона относительно вертикальной/крена) плюс 3 направления $x/y/z$ поступательного перемещения в пространстве.

Работа авторов Potard и др. расширяет понятие протяженности источника в качестве одномерного параметра источника (т.е. его ширины между двумя громкоговорителями) посредством изучения восприятия форм источников. В ней формируются несколько некогерентных точечных источников посредством применения (варьирующихся во времени) технологий декорреляции к первоначальному сигналу источника и затем помещения некогерентных источников в различные пространственные местоположения и за счет этого обеспечения им трехмерной протяженности.

В усовершенствованном стандарте AudioBIFS MPEG-4 объемные объекты/формы (раковина, коробка, эллипсоид и цилиндр) могут быть заполнены несколькими одинаково распределенными и декоррелированными источниками звука, чтобы обеспечить трехмерную протяженность источника.

Чтобы увеличивать и управлять протяженностью источника с использованием амбиофонии, в работе авторов Schmele и др. предложено смещение уменьшения порядка амбиофонии входного сигнала, что внутренне увеличивает кажущуюся ширину источника и распределение декоррелированных копий сигнала источника вокруг пространства для прослушивания.

Другой подход введен в работе авторов Zotter и др., в которой приспособливается принцип, предложенный в (т.е. извлечение пар фильтров, которые вводят частотно-зависимые разности фаз и абсолютных величин, чтобы достигать протяженности источника в компоновках для стереовоспроизведения) для амбиофонии.

Общий недостаток подходов на основе панорамирования (например) заключается в их зависимости от положения слушателя. Даже небольшое отклонение от зоны наилучшего восприятия приводит к тому, что пространственное изображение сворачивается до громкоговорителя, ближайшего к слушателю. Это радикально ограничивает их применение в контексте виртуальной реальности и дополненной реальности с 6 степенями свободы (6DoF), когда предполагается, что слушатель свободно перемещается. Кроме того, распределение частотно-временных элементов разрешения в подходах на основе DirAC (например) не всегда гарантирует надлежащий рендеринг пространственной протяженности фантомных источников. Кроме того, оно типично значительно ухудшает тембр сигнала источника.

Декорреляция сигналов источников обычно достигается посредством одного из следующих способов: i) извлечение пар фильтров с комплементарной абсолютной величиной (например), ii) использование всечастотных фильтров с постоянной абсолютной величиной, но (случайно) скремблированной фазой (например), или iii) пространственно случайное распределение частотно-временных элементов разрешения сигнала источника (например).

Все подходы влекут за собой свои последствия: Комплементарная фильтрация сигнала источника согласно i) типично приводит к измененному воспринимаемому тембру декоррелированных сигналов. Хотя всечастотная фильтрация, как указано в ii), сохраняет тембр сигнала источника, скремблированная фаза нарушает исходные соотношения фаз и, в частности, для переходных сигналов, вызывает серьезную временную дисперсию и артефакты размывания. Пространственное распределение частотно-временных элементов разрешения оказывается эффективным для некоторых сигналов, но также и изменяет воспринимаемый тембр сигнала. Кроме того, оно демонстрирует сильную зависимость от сигнала и вводит серьезные артефакты для импульсных сигналов.

Заполнение объемных форм несколькими декоррелированными версиями сигнала источника, предложенное в усовершенствованном AudioBIFS-стандарте, предполагает

доступность большого числа фильтров, которые формируют взаимно декоррелированные выходные сигналы (типично, более десяти точечных источников в расчете на объемную форму используются). Тем не менее, нахождение таких фильтров не представляет собой тривиальную задачу и становится тем более сложным, чем
5 больше таких фильтров требуется. Кроме того, если сигналы источников не полностью декоррелируются, и слушатель перемещается вокруг такой формы, например, в сценарии (виртуальной реальности), отдельные расстояния от источника до слушателя соответствуют различным задержкам сигналов источников, и их наложение в ушах слушателя приводит к зависимой от положения гребенчатой фильтрации, потенциально
10 вводящей раздражающее неустановившееся окрашивание сигнала источника.

Управление шириной источника с помощью технологию на основе амбиофонии в посредством понижения порядка амбиофонии демонстрирует слышимый эффект только для переходов от второго к первому или к нулевому порядку. Кроме того, эти переходы воспринимаются не только в качестве расширения источника, но также и часто в
15 качестве перемещения фантомного источника. При том, что суммирование декоррелированных версий сигнала источника может помогать в стабилизации восприятия кажущейся ширины источника, оно также вводит эффекты гребенчатой фильтрации, которые изменяют тембр фантомного источника.

Эффективный способ бинаурального рендеринга пространственно протяженного источника звука (SESS) раскрыт в WO2021/180935 с использованием двух
20 декоррелированных версий входной волновой формы сигнала (они могут формироваться посредством использования исходного моносигнала и декоррелятора, чтобы формировать декоррелированную версию этого моносигнала), каскада вычисления сигнальных меток, который вычисляет целевые бинауральные (и тембральные)
25 сигнальные метки пространственно протяженного источника звука в зависимости от размера источника (например, заданного в качестве диапазона углов азимута/подъема в зависимости от положения и ориентации пространственно протяженного источника звука и слушателя). В предпочтительном варианте осуществления, этот каскад вычисления сигнальных меток предварительно вычисляет целевые сигнальные метки
30 в зависимости от пространственных областей, которые должны покрываться посредством SESS, и сохраняет их в таблицу поиска, и каскад регулирования бинауральных сигнальных меток формирует подготовленный посредством бинаурального рендеринга выходной сигнал из входного сигнала, и его декоррелированная версия с использованием целевых сигнальных меток формирует
35 каскад вычисления сигнальных меток (таблицу поиска). Каскад регулирования бинауральных сигнальных меток регулирует бинауральные сигнальные метки (межканальную когерентность (ICC), межканальную разность фаз (ICPD), межканальную разность уровней (ICLD)) входных сигналов за несколько этапов до их требуемого целевого значения, вычисленного посредством каскада вычисления сигнальных меток/
40 таблицы поиска.

Задача настоящего изобретения состоит в создании усовершенствованной концепции для пространственно протяженных источников звука.

Данная задача решается объектами изобретения, определёнными в независимых пунктах формулы изобретения, и предпочтительные варианты осуществления
45 определены в зависимых пунктах формулы изобретения.

Регулярный алгоритм быстрого синтеза пространственно протяженных источников звука (SESS) имитирует звуковое впечатление рассеянного поля в определенных указанных целевых пространственных областях. Это достигается посредством

(виртуального) суммирования множества близкорасположенных источников звука, которые возбуждаются посредством декоррелированных версий аудиосигнала. Иногда, часть SESS загорается посредством частично пропускающего материала (например, кустарников), приводя к частотно-избирательному ослаблению SESS в загороженной 5 пространственной области. Этот эффект может разумно и эффективно включаться в эффективный алгоритм SESS посредством введения этапа взвешивания в вычисление между операцией табличного поиска и дополнительным вычислением требуемых бинауральных сигнальных меток. Таблица поиска сохраняет предварительно вычисленные частичные суммы членов для каждого пространственного сектора вокруг 10 слушателя. Протягивание достигается практически без дополнительных вычислительных затрат. Варианты осуществления относятся к устройству и к способу либо к компьютерной программе для воспроизведения или синтеза пространственно протяженного источника звука (SESS) с избирательным пространственным взвешиванием.

15 Преимущество настоящего изобретения заключается в том, что настоящее изобретение обеспечивает возможность обработки пространственно протяженного источника звука с возможно комплексной геометрической формой.

Дополнительное преимущество настоящего изобретения заключается в том, что варианты осуществления обеспечивают возможность усовершенствованной концепции 20 воспроизведения пространственно протяженного источника звука и обеспечивают возможности для пространственно избирательной модификации рендеринга SESS.

Первый аспект относится к использованию элементарных пространственных секторов. Этот первый аспект относится к сохранению данных для элементарных пространственных секторов в таблице поиска, при этом элементарные пространственные 25 секторы распределяются по сфере. Данные для элементарных пространственных секторов предпочтительно привязываются к голове пользователя, формирующей центрированную на пользователе аудиосцену, и являются одинаковыми для каждого наклона головы в одном и том же положении, а также для каждого положения головы слушателя, т.е. для каждой степени свободы из 6DOF. Тем не менее, каждое перемещение или наклон головы приводит к такой ситуации, когда звук из SESS «входит» в других 30 одном или более элементарных пространственных секторов в голову пользователя. Модуль рендеринга определяет элементарные пространственные секторы, покрытые посредством SESS, извлекает сохраненные данные для этих конкретных секторов, при необходимости выполняет взвешивание сохраненных данных вследствие 35 загорающих объектов или определенных расстояний, а затем комбинирует сохраненные данные (или в случае взвешивания взвешенных сохраненных данных) и после этого использует результат комбинированной операции для рендеринга (например, сигнальные метки рендеринга вычисляются из комбинированных данных (ко)вариации, но здесь также могут использоваться другие этапы и параметры). Следовательно, этот 40 аспект может использовать или может не использовать обращение к загорающим объектам и может использовать или может не использовать обращение к конкретным сохраненным данным дисперсии, поскольку комбинирование (и при необходимости также взвешивание) также может осуществляться, когда сохраняются другие данные, например, (средние) HRTF (для элементарного пространственного сектора или для 45 целой пространственной протяженности) либо даже непосредственно частотно-зависимые сигнальные метки.

Второй аспект относится к модифицирующим объектам, которые могут представлять собой загорающие объекты или другие объекты, приводящие к модификации звука

SESS на пути из положения SESS к пользователю, имеющему определенное местоположение и/или наклон. Этот второй аспект относится к обработке, например, загораживающих объектов. Воздействие загораживающего объекта представляет собой частотно-зависимое ослабление, имеющее характеристику нижних частот. Частотно-зависимое взвешивание также может применяться к процедуре из уровня техники, в которой отсутствуют элементарные пространственные секторы. На основе передаваемых данных, описывающих загораживающие объекты, следует принимать решение в отношении того, загораживается ли SESS, и затем применять функцию загораживания, например, к частотно-зависимым сохраненным сигнальным меткам, которые уже приводятся для различных частот в уровне техники. Следовательно, это представляет собой полезное применение эффекта загораживания в уровне техники без использования элементарных пространственных секторов или без использования сохраненных данных дисперсии.

Третий аспект относится к устройству хранения данных дисперсии и данных ковариации, например, для HRTF для различной пространственной протяженности или элементарных пространственных секторов. Этот третий аспект относится к устройству хранения, например, в таблице поиска, данных дисперсии и данных ковариации, например, для HRTF в положении хранения. Не имеет значения, сохраняются эти данные для определенной пространственной протяженности аналогично уровню техники или для элементарного пространственного сектора. Модуль рендеринга затем вычисляет все сигнальные метки рендеринга из сохраненных данных дисперсии на лету. В отличие от варианта применения из уровня техники, в котором сохраняется по меньшей мере IACC и вероятно другие сигнальные метки или данные HRFT, это не осуществляется в этом аспекте. Данные ковариации сохраняются, и сигнальные метки вычисляются на лету. Следовательно, этот аспект может использовать или может не использовать элементарные пространственные секторы и может использовать или может не использовать модифицирующие или загораживающие объекты.

Все аспекты могут использоваться отдельно друг от друга или вместе друг с другом, либо также только произвольно выбранные два аспекта могут комбинироваться.

Ниже описаны предпочтительные варианты осуществления настоящего изобретения с обращением к сопровождающим чертежам, на которых:

Фиг. 1 иллюстрирует устройство для синтезирования пространственно протяженного источника звука в соответствии с первым аспектом настоящего изобретения;

Фиг. 2a иллюстрирует устройство для синтезирования пространственно протяженного источника звука в соответствии со вторым аспектом изобретения;

Фиг. 2b иллюстрирует генератор аудиосцен в соответствии со вторым аспектом настоящего изобретения;

Фиг. 3 иллюстрирует предпочтительный вариант осуществления третьего аспекта настоящего изобретения;

Фиг. 4 иллюстрирует блок-схему для иллюстрации определенных частей аспектов согласно изобретению;

Фиг. 5 иллюстрирует другую блок-схему для иллюстрации нескольких частей аспектов согласно изобретению;

Фиг. 6 иллюстрирует дополнительную блок-схему для иллюстрации частей аспектов согласно изобретению;

Фиг. 7 иллюстрирует примерное разделение диапазона рендеринга на элементарные пространственные секторы;

Фиг. 8 иллюстрирует процедуру для комбинирования трех аспектов согласно

изобретению для синтеза пространственно протяженных источников звука;

Фиг. 9 иллюстрирует предпочтительную реализацию блока 320 по фиг. 4, 5 и 6;

Фиг. 10 иллюстрирует реализацию второго процессора каналов;

Фиг. 11 иллюстрирует принципиальную схему, подробно показывающую признаки
5 первого аспекта и второго аспекта изобретения;

На Фиг. 12 приведена иллюстрация для пояснения первого, второго и третьего аспектов согласно изобретению; и

Фиг. 13 иллюстрирует декоррелятор по фиг. 10, соединенный с синтезом аудиопроцессора в соответствии с дополнительным вариантом осуществления.

10 Фиг. 1 иллюстрирует устройство для синтезирования пространственно протяженного источника звука. Устройство содержит устройство 2000 хранения данных для сохранения элементов данных рендеринга для различных элементарных пространственных секторов, покрывающих диапазон рендеринга для слушателя. Кроме того, устройство содержит процессор 4000 идентификации секторов для идентификации, из различных элементарных
15 пространственных секторов, набора элементарных пространственных секторов, принадлежащих конкретному пространственно протяженному источнику звука. Идентификация выполняется на основе данных слушателя и данных, связанных с пространственно протяженным источником звука (SESS). Кроме того, устройство содержит модуль 5000 вычисления целевых данных для вычисления целевых данных
20 рендеринга из элементов данных рендеринга для набора элементарных пространственных секторов. Кроме того, устройство содержит аудиопроцессор 3000 для обработки аудиосигнала, представляющего пространственно протяженный источник звука, с использованием целевых данных рендеринга, сформированных посредством модуля 5000 вычисления целевых данных.

25 Фиг. 2а иллюстрирует устройство для синтезирования пространственно протяженного источника звука (SESS), содержащее входной интерфейс 4020 для приема описания аудиосцены, причем описание аудиосцены содержит данные пространственно протяженных источников звука по пространственно протяженному источнику звука и данные для модификации по потенциально модифицирующему объекту. Кроме того,
30 входной интерфейс 4020 выполнен с возможностью приема данных слушателя.

Процессор 4000 идентификации секторов, который, в общем, может реализовываться как процессор 4000 идентификации секторов по фиг. 1, выполнен с возможностью идентификации ограниченного модифицированного пространственного сектора для пространственно протяженного источника звука в диапазоне рендеринга для слушателя,
35 при этом диапазон рендеринга для слушателя больше ограниченного модифицированного пространственного сектора. Идентификация выполняется на основе данных пространственно протяженных источников звука и данных слушателя, и данных для модификации. Кроме того, устройство содержит модуль 5000 вычисления целевых данных, который, в общем, может быть реализован так же или реализован
40 аналогично модулю 5000 вычисления целевых данных по фиг. 1. Это устройство выполнено с возможностью вычисления целевых данных рендеринга из одного или более элементов данных рендеринга, принадлежащих модифицированному ограниченному пространственному сектору, определенному посредством блока 4000 по фиг. 2а. Кроме того, устройство для синтезирования пространственно протяженного
45 источника звука в соответствии со вторым аспектом, проиллюстрированным на фиг. 2а, содержит аудиопроцессор для обработки аудиосигнала, представляющего пространственно протяженный источник звука, с использованием целевых данных рендеринга, на которые влияют данные для модификации, т.е. данные по

модифицирующему объекту, например, загораживающему объекту.

Фиг. 2b иллюстрирует, снова в соответствии со вторым аспектом, генератор аудиосцен, содержащий генератор 6010 данных пространственно протяженных источников звука, генератор 6020 данных для модификации и выходной интерфейс 6030. Генератор 6010 данных пространственно протяженных источников звука выполнен с возможностью формирования данных пространственно протяженного источника звука и передачи этих данных в выходной интерфейс. Эти данные предпочтительно содержат по меньшей мере одно из информации местоположения и информации ориентации и геометрических данных для пространственно протяженного источника звука в качестве метаданных для пространственно протяженного источника звука и, кроме того, могут содержать данные формы сигнала для SESS, такие как стереосигнал для SESS в случае, например, большого SESS, такого как рояль, либо только моносигнал для данных SESS, которые обрабатываются посредством проиллюстрированного декоррелятора, например, на фиг. 10 в элементе 310 или на фиг. 13 в элементе 3100.

Генератор 6020 данных для модификации выполнен с возможностью формирования данных для модификации, и эти данные для модификации могут содержать описание функции нижних частот или описание геометрических данных относительно потенциально модифицирующего объекта. В варианте осуществления, функция нижних частот содержит значение ослабления для более высокой частоты, причем значение ослабления для более высокой частоты, представляющее значение ослабления, является более сильным по сравнению со значением ослабления для более низкой частоты, и эти данные перенаправляются в выходной интерфейс 6030 для вставки в сформированное описание аудиосцен.

Следовательно, описание аудиосцен, проиллюстрированное на фиг. 2b, улучшается по сравнению с описанием SESS в том, что включаются не только данные SESS, но также и данные по объектам для модификации, которые, сами по себе, представляют собой не источники звука, а которые представляют собой элементы, которые модифицируют звуковое поле, сформированное посредством источника звука.

Фиг. 3 иллюстрирует предпочтительный вариант осуществления устройства для синтеза пространственно протяженного источника звука в соответствии с третьим аспектом.

Этот элемент содержит устройство хранения данных для сохранения одного или более элементов данных рендеринга для различных ограниченных пространственных секторов, при этом различные ограниченные пространственные секторы расположены в диапазоне рендеринга для слушателя, и при этом один или более элементов данных рендеринга для ограниченного пространственного сектора содержат по меньшей мере один из левого элемента данных дисперсии, правого элемента данных дисперсии и лево-правого элемента данных ковариации.

Кроме того, устройство содержит процессор 4000 идентификации секторов для идентификации одного или более ограниченных пространственных секторов для пространственно протяженного источника звука в диапазоне рендеринга для слушателя на основе данных пространственно протяженных источников звука и предпочтительно на основе положения или ориентации слушателя.

Левые данные дисперсии, правые данные дисперсии и данные ковариации вводятся в модуль 5000 вычисления целевых данных для вычисления целевых данных рендеринга из сохраненных левых данных дисперсии, сохраненных правых данных дисперсии или сохраненных данных ковариации, соответствующих одному или более ограниченных пространственных секторов, определенных посредством процессора 4000 идентификации

секторов. Целевые данные рендеринга перенаправляются в аудиопроцессор 3000 для обработки аудиосигнала, представляющего пространственно протяженный источник звука, с использованием целевых данных рендеринга. В общем, аудиопроцессор 3000 может реализовываться аналогично тому, что показано на фиг. 1 и 2b или фиг. 4, 5 и 6, либо аудиопроцессор 3000 может реализовываться по-другому.

Предпочтительно, левый элемент данных дисперсии, правый элемент данных дисперсии и/или лево-правые элементы данных ковариации представляют собой элементы данных, связанные с данными передаточной функции восприятия звука человеком или связанные с данными бинаурального импульсного отклика в помещении, или связанные с данными бинауральной передаточной функции в помещении, или связанные с данными импульсного отклика восприятия звука человеком. Кроме того, элементы данных рендеринга содержат значения элементов данных дисперсии или ковариации для различных частот таким образом, что частотно-избирательная/частотно-зависимая обработка достигается.

В частности, устройство 2000 хранения данных выполнено с возможностью сохранения, для каждого ограниченного пространственного сектора, частотно-зависимого представления левого элемента данных дисперсии, частотно-зависимого представления правого элемента данных дисперсии и частотно-зависимого представления элемента данных ковариации.

Восходящая обработка сохраненных элементов данных дисперсии/ковариации примерно иллюстрируется на нескольких чертежах из WO2021/180935, указываемых далее как фиг. 4, 5 и 6.

Фиг. 4 показывает блок-схему синтеза SESS. Фиг. 5 показывает другую блок-схему синтеза SESS, упрощенного в соответствии с вариантом 1, и фиг. 6 показывает блок-схему синтеза SESS, упрощенного в соответствии с вариантом 2.

Фиг. 4 иллюстрирует реализацию устройства для синтезирования пространственно протяженного источника звука. Устройство содержит интерфейс обработки пространственной информации, который принимает ввод информации индикатора пространственного диапазона, указывающий ограниченный пространственный диапазон для пространственно протяженного источника звука в пределах максимального пространственного диапазона. Ограниченный пространственный диапазон вводится в блок 200 обеспечения информации сигнальных меток, выполненный с возможностью обеспечения одного или более информационных элементов сигнальных меток в ответ на ограниченный пространственный диапазон, обеспеченный интерфейсом обработки пространственной информации. Информационный элемент сигнальных меток либо несколько информационных элементов сигнальных меток вводятся в аудиопроцессор 300, выполненный с возможностью обработки аудиосигнала, представляющего пространственно протяженный источник звука, с использованием одного или более информационных элементов сигнальных меток, обеспеченных блоком 200 обеспечения информации сигнальных меток. Аудиосигнал для пространственно протяженного источника звука (SESS) может представлять собой один канал или может представлять собой первый аудиоканал и второй аудиоканал, или может представлять собой более двух аудиоканалов. Тем не менее, для целей наличия низкой нагрузки по обработке, небольшое число каналов для пространственно протяженного источника звука или для аудиосигнала, представляющего пространственно протяженный источник звука, является предпочтительным.

Аудиосигнал вводится в аудиопроцессор 300, и аудиопроцессор 300 обрабатывает входной аудиосигнал, либо, когда число входных аудиоканалов меньше, чем требуется,

например, составляет только один, аудиопроцессор содержит второй процессор 310 каналов, проиллюстрированный на фиг. 10, содержащий, например, декоррелятор для формирования второго аудиоканала S_2 , декоррелированного относительно первого аудиоканала S , который также проиллюстрирован на фиг. 10 в качестве S_1 .

5 Информационные элементы сигнальных меток могут представлять собой фактические элементы сигнальных меток, такие как элементы межканальной корреляции, элементы межканальной разности фаз, элементы межканальной разности уровней и усиления, элементы G_1 , G_2 коэффициентов усиления, совместно представляющие межканальную
10 разность уровней и/или абсолютные уровни амплитуды или мощности, или энергии, например, либо информационные элементы сигнальных меток также могут представлять собой фактические функции фильтра, такие как передаточные функции восприятия звука человеком с числом, требуемым посредством фактического числа подлежащих синтезированию выходных каналов в синтезирующем сигнале. Таким образом, когда
15 синтезирующий сигнал должен иметь два канала, например, два бинауральных канала или два канала громкоговорителей, одна передаточная функция восприятия звука человеком для каждого канала требуется. Вместо передаточных функций восприятия звука человеком, необходимы функции импульсного отклика восприятия звука человеком (HRIR) или функции бинаурального или небинаурального импульсного
20 отклика в помещении ((B)RIR). Одна такая передаточная функция требуется для каждого канала, и фиг. 4 иллюстрирует реализацию наличия двух каналов.

В варианте осуществления, блок 200 обеспечения информации сигнальных меток выполнен с возможностью обеспечения в качестве информационного элемента
25 сигнальных меток значения межканальной корреляции. Аудиопроцессор 300 выполнен с возможностью фактического приёма через аудиосигнальный интерфейс 305 первого аудиоканала и второго аудиоканала. Тем не менее, когда аудиосигнальный интерфейс 305 принимает только один канал, при необходимости предусмотренный второй процессор каналов формирует, например, посредством процедуры на фиг. 9, второй аудиоканал. Аудиопроцессор выполняет обработку корреляции, чтобы налагать
30 корреляцию между первым аудиоканалом и вторым аудиоканалом с использованием значения межканальной корреляции.

В качестве альтернативы или дополнения, может быть предусмотрен дополнительный информационный элемент сигнальных меток, такой как элемент межканальной разности фаз, элемент межканальной разности времен, элемент межканальной разности уровней
35 и усиления либо информационный элемент первых коэффициентов усиления и вторых коэффициентов усиления. Элементы также могут представлять собой значения интерауральной корреляции (IACC), т.е. более конкретные значения межканальной корреляции, или элементы интерауральной разности фаз (IAPD), т.е. более конкретные значения межканальной разности фаз.

40 В предпочтительном варианте осуществления, корреляция налагается (320) посредством аудиопроцессора 300 в ответ на информационный элемент сигнальных меток корреляции, до того, как выполняются регулирование ICPD (330), ICTD или ICLD (340), либо до того, как выполняются HRTF или другие функциональные обработки (350) передаточного фильтра. Тем не менее, в зависимости от обстоятельств, порядок
45 может задаваться по-другому.

В предпочтительном варианте осуществления, устройство содержит запоминающее устройство для сохранения информации касательно различных информационных элементов сигнальных меток относительно различных индикаторов пространственного

диапазона. В этой ситуации блок обеспечения информации сигнальных меток дополнительно содержит выходной интерфейс для извлечения, из запоминающего устройства, одного или более информационных элементов сигнальных меток, ассоциированных с индикатором пространственного диапазона, вводимым в соответствующее запоминающее устройство. Такая таблица 210 поиска, например, иллюстрируется на фиг. 4, 5 или 6, причем таблица поиска содержит запоминающее устройство и выходной интерфейс для вывода соответствующих информационных элементов сигнальных меток. В частности, запоминающее устройство может не только сохранять значения IACC, IAPD или G_l и G_r , как проиллюстрировано на фиг. 1b, но запоминающее устройство в таблице поиска также может сохранять функции фильтра, как проиллюстрировано в блоке 220 по фиг. 5 и по фиг. 6, что указано как «выбор HRTF». В этом варианте осуществления, хотя они проиллюстрированы отдельно на фиг. 5 и на фиг. 6, блоки 210, 220 могут содержать одно и то же запоминающее устройство, причем, в ассоциации с соответствующим индикатором пространственного диапазона, указываемым в качестве углов азимута и углов подъема, соответствующие информационные элементы сигнальных меток, такие как IACC и, при необходимости, IAPD, и передаточные функции для фильтров, такие как $HRTF_l$ для левого выходного канала и $HRTF_r$ для правого выходного канала сохраняются, при этом левый и правый выходные каналы указываются в качестве S_l и S_r на фиг. 4 или на фиг. 5, или на фиг. 6.

Запоминающее устройство, используемое посредством таблицы 210 поиска или блока 220 выбора функции, также может использовать устройство хранения данных, в котором на основе определенных кодов секторов либо углов секторов или диапазонов углов секторов, доступны соответствующие параметры. Альтернативно, запоминающее устройство может сохранять процедуру обработки векторной таблицы кодирования или многомерной подгонки функции либо гауссову смешанную модель (GMM) или метод опорных векторов (SVM) в зависимости от обстоятельств.

Целевые сигнальные метки вычисляются так, как описано ниже по тексту. На фиг. 4, показывается общая блок-схема концепции. $[\Phi_1, \Phi_2]$ описывает требуемую протяженность источника с точки зрения диапазона углов азимута. $[\theta_1, \theta_2]$ является требуемой протяженностью источника с точки зрения диапазона углов подъема. $S_1(\omega)$ и $S_2(\omega)$ обозначают два декоррелированных входных сигнала, где ω описывает частотный индекс. Таким образом, для $S_1(\omega)$ и $S_2(\omega)$, следующее уравнение является справедливым:

$$E\{S_1(\omega) \cdot S_2^*(\omega)\} = 0. \quad (1)$$

Кроме того, оба входных сигнала должны иметь равную спектральную плотность мощности. В качестве альтернативы, можно предусмотреть только один входной сигнал, $S(\omega)$. Второй входной сигнал формируется внутренне с использованием декоррелятора, как проиллюстрировано на фиг. 10. При условии $S_l(\omega)$ и $S_r(\omega)$, протяженный источник звука синтезируется посредством последовательного регулирования межканальной когерентности (ICC), межканальных разностей фаз (ICPD) и межканальных разностей уровней (ICLD) таким образом, что они совпадают с соответствующими интерауральными сигнальными метками. Величины, необходимые для этих этапов обработки, считываются из предварительно вычисленной таблицы поиска. Результирующие сигналы левого и правого канала, $S_l(\omega)$ и $S_r(\omega)$, могут воспроизводиться через наушники и напоминать SESS. Следует отметить, что регулирование ICC должно

выполняться сначала; тем не менее, блоки регулирования ICPD и ICLD могут меняться местами. Вместо IAPD также могут воспроизводиться соответствующие интерауральные разности времен (IATD). Тем не менее, далее подробнее рассматривается только IAPD.

В блоке регулирования ICC, взаимная корреляция между обоими входными сигналами регулируется до требуемого значения $|IACC(\omega)|$ с использованием следующих формул [21]:

$$\hat{S}_1(\omega) = H_\alpha(\omega) \cdot S_1(\omega) + H_\beta(\omega) \cdot S_2(\omega), \quad (2)$$

$$\hat{S}_2(\omega) = H_\alpha(\omega) \cdot S_2(\omega) + H_\beta(\omega) \cdot S_1(\omega), \quad (3)$$

$$H_\beta(\omega) = H_\alpha(\omega) \sqrt{\frac{1}{2}(1 - \sqrt{1 - |IACC(\omega)|^2})}, \quad (4)$$

$$H_\alpha(\omega) = \sqrt{1 - H_\beta^2(\omega)}. \quad (5)$$

Применение этих формул приводит к требуемой взаимной корреляции, при условии, что входные сигналы $S_1(\omega)$ и $S_2(\omega)$ полностью декоррелируются. Кроме того, их спектральная плотность мощности должна быть равной. Соответствующая блок-схема показывается на фиг. 9. Четыре фильтра 321-324 и два сумматора 325, 326 обрабатывают ввод, чтобы получить вывод блока 320.

Блок 330 регулирования ICPD описан посредством следующих формул:

$$\hat{S}'_1(\omega) = e^{j \cdot IAPD(\omega)} \cdot \hat{S}_1(\omega), \quad (6)$$

$$\hat{S}'_2(\omega) = \hat{S}_2(\omega). \quad (7)$$

В завершение, регулирование 340 ICLD выполняется следующим образом:

$$S_l(\omega) = G_l(\omega) \cdot \hat{S}'_1(\omega), \quad (8)$$

$$S_r(\omega) = G_r(\omega) \cdot \hat{S}'_2(\omega), \quad (9)$$

где $G_l(\omega)$ описывает усиление для левого уха, и $G_r(\omega)$ описывает усиление для правого уха. Это приводит к требуемой ICLD при условии, что $\hat{S}'_1(\omega)$ и $\hat{S}'_2(\omega)$ фактически имеют равную спектральную плотность мощности. Поскольку усиление для левого и правого уха используется непосредственно, моноауральные спектральные сигнальные метки воспроизводятся в дополнение к IALD.

Чтобы дополнительно упростить поясняемый выше способ, описаны два варианта упрощения. Как упомянуто выше, основная интерауральная сигнальная метка, оказывающая влияние на воспринимаемую пространственную протяженность (в горизонтальной плоскости), представляет собой IACC. В силу этого должна быть возможность не использовать предварительно вычисленные IAPD и/или значения IALD, а регулировать значения IAPD и/или IALD непосредственно через HRTF. С этой целью, используется HRTF, соответствующая положению, представляющему требуемый диапазон протяженностей источников. В качестве этого положения, среднее требуемого диапазона изменения азимута/подъема выбирается здесь без потери общности. Ниже по тексту приводится описание обоих вариантов.

Первый вариант включает в себе использование предварительно вычисленных

IACC- и IAPD-значений. Тем не менее, ICLD регулируется с использованием HRTF, соответствующей центру диапазона протяженностей источников.

Блок-схема первого варианта показывается на фиг. 5. $S_l(\omega)$ и $S_r(\omega)$ теперь вычисляются с использованием следующих формул:

$$S_l(\omega) = \hat{S}'_1(\omega) \cdot |\text{HRTF}_l(\omega, \bar{\phi}, \bar{\theta})|, \quad (10)$$

$$S_r(\omega) = \hat{S}'_2(\omega) \cdot |\text{HRTF}_r(\omega, \bar{\phi}, \bar{\theta})|, \quad (11)$$

где $\bar{\phi} = (\phi_1 + \phi_2)/2$ и $\bar{\theta} = (\theta_1 + \theta_2)/2$ описывают местоположение HRTF, которая представляет среднее требуемого диапазона изменения азимута/подъема. Основные преимущества первого варианта включают в себя:

- Отсутствие формирования спектра/окрашивания, когда протяженность источника увеличивается по сравнению с точечным источником в центре диапазона протяженностей источников.

- Более низкие требования к запоминающему устройству по сравнению с полнофункциональным способом, поскольку $G_l(\omega)$ и $G_r(\omega)$ не должны обязательно сохраняться в таблице поиска.

- Большая гибкость к изменениям набора данных HRTF во время выполнения по сравнению с полнофункциональным способом, поскольку только результирующие ICC и ICPD, а не ICLD, зависят от набора данных HRTF, используемого во время предварительного вычисления.

Основной недостаток этой упрощенной версии заключается в том, что она должна сбиться каждый раз, когда радикальные изменения в IALD возникают, по сравнению с непротяженным источником. В этом случае, IALD не должна воспроизводиться с достаточной точностью. Это, например, имеет место, когда источник не центрируется вокруг азимута в 0° , и в то же время, протяженность источника в горизонтальном направлении становится слишком большой.

Второй вариант включает в себе использование только предварительно вычисленных IACC-значений. ICPD и ICLD регулируются с использованием HRTF, соответствующей центру диапазона протяженностей источников.

Блок-схема второго варианта показывается на фиг. 6. $S_l(\omega)$ и $S_r(\omega)$ теперь вычисляются с использованием следующих формул:

$$S_l(\omega) = \hat{S}_1(\omega) \cdot \text{HRTF}_l(\omega, \bar{\phi}, \bar{\theta}), \quad (12)$$

$$S_r(\omega) = \hat{S}_2(\omega) \cdot \text{HRTF}_r(\omega, \bar{\phi}, \bar{\theta}). \quad (13)$$

В отличие от первого варианта, фаза и абсолютная величина HRTF теперь используются вместо только абсолютной величины. Это обеспечивает возможность регулировать не только ICLD, но также и ICPD.

Во-первых, (ко)-вариационные члены вычисляются между левым и правым каналом следующим образом:

$\mathbf{E}\{Y_l(\omega) \cdot Y_r^*(\omega)\}$, $\mathbf{E}\{|Y_l(\omega)|^2\}$ и $\mathbf{E}\{|Y_r(\omega)|^2\}$ извлекаются:

$$\begin{aligned}
 \mathbf{E}\{Y_l(\omega) \cdot Y_r^*(\omega)\} &= \mathbf{E} \left\{ \sum_{n=1}^N A_{l,n} \cdot e^{j\phi_{l,n}} \cdot S_n(\omega) \cdot \sum_{m=1}^N A_{r,m} \cdot e^{-j\phi_{r,m}} \cdot S_m^*(\omega) \right\} \\
 &= \mathbf{E} \left\{ \sum_{n=1}^N \sum_{m=1}^N A_{l,n} \cdot A_{r,m} \cdot e^{j(\phi_{l,n} - \phi_{r,m})} \cdot S_n(\omega) \cdot S_m^*(\omega) \right\} \\
 &= \sum_{n=1}^N \sum_{m=1}^N A_{l,n} \cdot A_{r,m} \cdot e^{j(\phi_{l,n} - \phi_{r,m})} \cdot \mathbf{E}\{S_n(\omega) \cdot S_m^*(\omega)\} \\
 &= P(\omega)^2 \cdot \sum_{n=1}^N A_{l,n} \cdot A_{r,n} \cdot e^{j(\phi_{l,n} - \phi_{r,n})},
 \end{aligned} \tag{20}$$

$$\begin{aligned}
 \mathbf{E}\{|Y_l(\omega)|^2\} &= \mathbf{E} \left\{ \sum_{n=1}^N A_{l,n} \cdot e^{j\phi_{l,n}} \cdot S_n(\omega) \cdot \sum_{m=1}^N A_{l,m} \cdot e^{-j\phi_{l,m}} \cdot S_m^*(\omega) \right\} \\
 &= \mathbf{E} \left\{ \sum_{n=1}^N \sum_{m=1}^N A_{l,n} \cdot A_{l,m} \cdot e^{j(\phi_{l,n} - \phi_{l,m})} \cdot S_n(\omega) \cdot S_m^*(\omega) \right\} \\
 &= \sum_{n=1}^N \sum_{m=1}^N A_{l,n} \cdot A_{l,m} \cdot e^{j(\phi_{l,n} - \phi_{l,m})} \cdot \mathbf{E}\{S_n(\omega) \cdot S_m^*(\omega)\} \\
 &= P(\omega)^2 \cdot \sum_{n=1}^N A_{l,n}^2,
 \end{aligned} \tag{21}$$

$$\mathbf{E}\{|Y_r(\omega)|^2\} = P(\omega)^2 \cdot \sum_{n=1}^N A_{r,n}^2. \tag{22}$$

На втором этапе, целевые сигнальные метки IACC, IALD и IAPD вычисляются из дисперсионных членов следующим образом:

$$\begin{aligned}
 \text{IACC}(\omega) &= \frac{\mathbf{E}\{Y_l(\omega) \cdot Y_r^*(\omega)\}}{\sqrt{\mathbf{E}\{|Y_l(\omega)|^2\} \cdot \mathbf{E}\{|Y_r(\omega)|^2\}}} \\
 &= \frac{\sum_{n=1}^N A_{l,n} \cdot A_{r,n} \cdot e^{j(\phi_{l,n} - \phi_{r,n})}}{\sqrt{\sum_{n=1}^N A_{l,n}^2 \cdot \sum_{m=1}^N A_{r,m}^2}},
 \end{aligned} \tag{23}$$

$$\begin{aligned}
 \text{IALD}(\omega) &= 10 \log_{10} \frac{\mathbf{E}\{|Y_l(\omega)|^2\}}{\mathbf{E}\{|Y_r(\omega)|^2\}} \\
 &= 10 \log_{10} \frac{\sum_{n=1}^N A_{l,n}^2}{\sum_{n=1}^N A_{r,n}^2},
 \end{aligned} \tag{24}$$

$$\begin{aligned}
 \text{IAPD}(\omega) &= \angle(\mathbf{E}\{Y_l(\omega) \cdot Y_r^*(\omega)\}) \\
 &= \angle(\text{IACC}(\omega)) \\
 &= \angle \left(\sum_{n=1}^N A_{l,n} \cdot A_{r,n} \cdot e^{j(\phi_{l,n} - \phi_{r,n})} \right).
 \end{aligned} \tag{25}$$

как и усиления для левого и правого уха:

$$G_l(\omega) = \sqrt{\frac{\mathbf{E}\{|Y_l(\omega)|^2\}}{N \cdot P(\omega)^2}} = \sqrt{\frac{\sum_{n=1}^N A_{l,n}^2}{N}} \quad (26)$$

$$G_r(\omega) = \sqrt{\frac{\mathbf{E}\{|Y_r(\omega)|^2\}}{N \cdot P(\omega)^2}} = \sqrt{\frac{\sum_{n=1}^N A_{r,n}^2}{N}} \quad (27)$$

Из этих целевых сигнальных меток, конечный эффективный синтез бинаурального сигнала может выполняться посредством проектирования 4 фильтров, преобразующих входной звук в рендерлируемый бинауральный вывод, как пояснено в WO2021/180935.

Первый аспект относится к использованию элементарных пространственных секторов. Этот первый аспект относится к сохранению данных для элементарных пространственных секторов в таблице поиска, при этом элементарные пространственные секторы распределяются по сфере. Данные для элементарных пространственных секторов предпочтительно привязываются к голове пользователя, формирующей

центрированную на пользователе аудиосцену, и являются одинаковыми для каждого наклона головы в одном и том же положении, а также для каждого положения головы слушателя, т.е. для каждой степени свободы из 6DOF. Тем не менее, каждое перемещение или наклон головы приводит к такой ситуации, когда звук из SESS «входит» в других одном или более элементарных пространственных секторов в голову пользователя. Модуль рендеринга определяет элементарные пространственные секторы, покрытые посредством SESS, извлекает сохраненные данные для этих конкретных секторов, при необходимости выполняет взвешивание сохраненных данных вследствие загромождающих объектов или определенных расстояний, а затем комбинирует сохраненные данные (или в случае взвешивания взвешенных сохраненных данных) и после этого использует результат комбинированной операции для рендеринга (например, сигнальные метки рендеринга вычисляются из комбинированных данных (ко)вариации, но здесь также могут использоваться другие этапы и параметры). Следовательно, этот аспект может использовать или может не использовать обращение к загромождающим объектам и может использовать или может не использовать обращение к конкретным сохраненным данным дисперсии, поскольку комбинирование (и при необходимости также взвешивание) также может осуществляться, когда сохраняются другие данные, например, (средние) HRTF (для элементарного пространственного сектора или для целой пространственной протяженности) либо даже непосредственно частотно-зависимые сигнальные метки.

Второй аспект относится к модифицирующим объектам, которые могут представлять собой загромождающие объекты или другие объекты, приводящие к модификации звука SESS на пути из положения SESS к пользователю, имеющему определенное местоположение и/или наклон. Этот второй аспект относится к обработке, например, загромождающих объектов. Воздействие загромождающего объекта представляет собой частотно-зависимое ослабление, имеющее характеристику нижних частот. Частотно-зависимое взвешивание также может применяться к процедуре из уровня техники, в которой отсутствуют элементарные пространственные секторы. На основе передаваемых данных, описывающих загромождающие объекты, следует принимать решение в отношении того, загромождается ли SESS, и затем применять функцию загромождания, например, к частотно-зависимым сохраненным сигнальным меткам, которые уже приводятся для различных частот в уровне техники. Следовательно, это представляет собой полезное применение эффекта загромождания в уровне техники без использования элементарных пространственных секторов или без использования

сохраненных данных дисперсии.

Третий аспект относится к устройству хранения данных дисперсии и данных ковариации, например, для HRTF для различной пространственной протяженности или элементарных пространственных секторов. Этот третий аспект относится к устройству хранения, например, в таблице поиска, данных дисперсии и данных ковариации, например, для HRTF в положении хранения. Не имеет значения, сохраняются эти данные для определенной пространственной протяженности аналогично уровню техники или для элементарного пространственного сектора. Модуль рендеринга затем вычисляет все сигнальные метки рендеринга из сохраненных данных дисперсии на лету. В отличие от варианта применения из уровня техники, в котором сохраняется по меньшей мере IACC и вероятно другие сигнальные метки или HRFT-данные, это не осуществляется в этом аспекте. Данные ковариации сохраняются, и сигнальные метки вычисляются на лету. Следовательно, этот аспект может использовать или может не использовать элементарные пространственные секторы и может использовать или может не использовать модифицирующие или загораживающие объекты.

Все аспекты могут использоваться отдельно друг от друга или вместе друг с другом, либо также только произвольно выбранные два аспекта могут комбинироваться.

Преимущество настоящего изобретения состоит в обеспечении усовершенствованного эффективного и реалистичного бинаурального рендеринга для пространственно протяженного источника звука по сравнению с WO2021/180935, например, посредством:

- организации таблицы поиска для вычисления целевых сигнальных меток конкретным способом (секторальным, с использованием (ко)-вариационных членов, частотно-зависимым); или

- выполнения (частотно-избирательного) взвешивания (ко)-вариационных членов согласно требуемому целевому частотному отклику, при необходимости для синтеза (частично или полностью) загороженных частей SESS либо для моделирования ослабления в зависимости от расстояния наверняка.

Варианты осуществления настоящего изобретения расширяют вышеописанную концепцию из WO2021/180935 для эффективного рендеринга SESS несколькими способами для повышения эффективности хранения и обеспечения характеристик рендеринга также частично загороженных частей SESS:

Раскрыт очень эффективный способ организации таблицы поиска и вычисления целевых сигнальных меток на основе таблицы поиска, который обеспечивает возможность покрывать все возможные пространственные целевые области для SESS в таблицу поиска с небольшим размером. Это достигается посредством организации таблицы поиска в качестве таблицы, которая сегментирует всю сферу вокруг головы слушателя на небольшие секторы азимута/подъема. Размер этих секторов (т.е. их размер азимута и подъема) предпочтительно выбирается в соответствии с разрешением человеческого восприятия азимута/подъема. Например, человеческое слуховое разрешение для азимута является самым точным (приблизительно 1 градус) впереди и снижается в направлении вбок. Кроме того, разрешение в восприятии подъема является гораздо более приблизительным, чем разрешение в азимуте вследствие того, что уши слушателя расположены слева и справа на голове. Для каждого из этих пространственных секторов, конкретные частично суммированные члены сохраняются в таблице поиска. В предпочтительном варианте осуществления, они представляют собой (ко)-вариационные члены ($E\{Y_l \cdot Y_r^*\}$, $E\{|Y_l|^2\}$, $E\{|Y_r|^2\}$) двух сигналов в уши, когда множество точечных источников (описанных посредством их соответствующих импульсных откликов восприятия звука человеком (HRIR) и возбужденных посредством

декоррелированных версий сигнала=рассеянное поле) суммируются. Кроме того, в предпочтительном варианте осуществления, эти записи таблицы сохраняются частотно-избирательным способом ($E\{Y_l \cdot Y_r^*}\}$, $E\{|Y_l|^2}\}$, $E\{|Y_r|^2}\}$).

5 Это также достигается отдельно или в дополнение к вышеуказанному, поскольку процесс вычисления сигнальных меток использует эти суммированные члены ($E\{Y_l \cdot Y_r^*}\}$, $E\{|Y_l|^2}\}$, $E\{|Y_r|^2}\}$) из HRIR-долей, которые сохраняются для каждого пространственного сектора, так что (когда должны покрываться несколько секторов) данные (ко)вариации для этих секторов могут просто суммироваться, чтобы формировать данные (ко) вариации для всей целевой области (включающей в себя все секторы).

10 Кроме того, пространственное взвешивание определенных пространственных секторов (например, чтобы моделировать загораживание этой части SESS) может достигаться посредством взвешивания данных (ко)вариации, сохраненных для этих пространственных секторов, перед их использованием в последующем процессе вычисления сигнальных меток. В частности, требуемый целевой частотный отклик $g(f)$ может налагаться посредством умножения всех (ко)-вариационных членов на соответствующий коэффициент $g^2(f)$ масштабирования энергии. В качестве примера, загораживающий кустарник должен налагать ослабление и частотный отклик нижних частот, когда звук распространяется через него. Таким образом, (ко)-вариационные члены должны ослабляться, и члены более высоких частот ослабляются больше членов низких частот. Возможны несколько зон для различных загораживаний/взвешивания. Аналогичным, также является возможным моделирование расстояния до объекта: Для больших объектов, таких как реки, части объекта могут находиться значительно дальше от слушателя, чем другие, в силу этого внося меньший уровень громкости, чем близлежащие части. Это может моделироваться и рендерироваться посредством взвешивания расстояния различных пространственных секторов. Члены в пространственных секторах взвешиваются с коэффициентом ослабления энергии в зависимости от расстояния, соответствующим (например, среднему) расстоянию объекта в этом пространственном секторе.

25 30 Ниже приведено общее представление варианта осуществления способа или устройства или компьютерной программы согласно изобретению.

В фазе инициализации/запуска модуля рендеринга, сегментация сферы вокруг головы слушателя осуществляется посредством задания пространственных секторов (например, диапазонов углов азимута и подъема), по которым впоследствии могут суммироваться доли HRIR. Затем, на основе этих пространственных секторов, соответствующие доли HRIR могут сохраняться в таблице поиска с использованием (ко)-вариационных членов.

35 40 Фиг. 11 иллюстрирует дополнительное общее представление настоящего изобретения (способа или устройства, или компьютерной программы), реализующее взаимодействие первого аспекта и второго аспекта. В частности, блок «выбор пространственных секторов для рендеринга SESS» соответствует процессору 4000 идентификации секторов, проиллюстрированному на фиг. 1-3. Результат выбора пространственных секторов представляет собой группу пространственных секторов, в которой могут быть предусмотрены некоторые секторы вообще без модификации, проиллюстрированной в 4010. Кроме того, из определенных секторов, могут быть предусмотрены секторы с модификацией/загораживанием в соответствии с первой характеристикой, проиллюстрированной в 4020. Кроме того, также могут быть предусмотрены секторы с другой модификацией загораживания, проиллюстрированной как «номер N». Это проиллюстрировано в 4030. Конкретное вычисление целевых данных,

проиллюстрированное посредством модуля 5000 вычисления целевых данных, в частности, для второго аспекта, выполняет суммирование дисперсионных членов для левой стороны, дисперсионных членов для правой стороны и ковариационных членов для всех незагороженных секторов в случае, если предусмотрено более одного такого сектора. Кроме того, суммирование в соответствии с функцией взвешивания 1 выполняется, т.е. если предусмотрено более 1 сектора с загораживанием в соответствии с модификацией/загораживанием номер 1, они суммируются, и затем соответствующий весовой коэффициент применяется, либо операция взвешивания и операция суммирования могут меняться местами. Кроме того, в случае если предусмотрены другие секторы с модификацией/загораживанием номер N, как проиллюстрировано в 4030, такие секторы могут суммироваться с соответствующим весовым коэффициентом для конкретной функции взвешивания/модификации для этих секторов.

Естественно, может возникать такой случай, в котором только незагороженные секторы являются существующими для SESS, либо имеются только загороженные секторы в соответствии с одной функцией модификации, либо любое сочетание этих возможностей, т.е. один незагороженный сектор и один сектор с модификацией/загораживанием номер 1, но ни одного для модификации/загораживания номер N. Естественно, номер N также может быть равен 1, так что существуют только строки 4010 и 4020, но любая модификация с другой модификацией поверх модификации номер 1 не определяется посредством блока 4000.

После того, как отдельное взвешивание для отдельных модификаций/загораживаний выполнено в блоке 5020, полное суммирование сигнальных меток в блоке 5040 осуществляется, и затем входные данные для конечного вычисления 5060 целевых сигнальных меток выполняются. Эти данные целевых сигнальных меток затем вводятся в блок 3000 синтеза бинауральных сигнальных меток или аудиопроцессора по фиг. 11. Ввод в блок 3000 представляет собой входной сигнал SESS номер 1 и входной сигнал SESS номер 2, если SESS имеет волновую форму стереосигнала. Тем не менее, в случае SESS, имеющего только волновую форму моносигнала, два сигнала формируются, но с помощью декоррелятора, проиллюстрированного в 3100 на фиг. 13 либо проиллюстрированного в 3010 на фиг. 10.

Фиг. 12 иллюстрирует предпочтительную реализацию синтеза 3000 бинауральных сигнальных меток, состоящего из регулирования 3200 IACC, регулирования 3300 IAPD и регулирования 3400 IALD. Все эти блоки содержат данные из устройства хранения данных, указываемого в качестве «таблицы поиска» в блоке 2000. Тем не менее, в зависимости от реализации, соответствующие обработки для определения конечных значений для IACC, IAPD и IALD также формируются в блоке 2000 в соответствии с этапами 5020, 5040, 5060 вычисления целевых данных. Следовательно, блок с названием «таблица поиска» на фиг. 12 приведён со ссылочной позицией 2000 и со ссылочной позицией 5000. Тем не менее, ввод в этот блок обеспечивается процессором 4000 идентификации секторов любого из фиг. 1, 2а, 3, 11.

Фиг. 13 иллюстрирует, в левой стороне, декоррелятор 3100 для формирования, из одной волновой формы сигнала SESS, двух входных сигналов SESS номер 1 и номер 2 в выводе декоррелятора. Эти данные затем подвергаются четырем операциям 3210, 3220, 3230 и 3240 фильтрации, при которых соответствующие доли для левого канала суммируются через сумматор 3250, и при которых соответствующие доли правого канала суммируются через сумматор 3260 для получения левого и правого конечных выходных сигналов. Отдельные функции 3210, 3220, 3230 и 3240 фильтра либо вычисляются через модуль 5000 вычисления целевых данных, соответственно, для

определенного ограниченного пространственного диапазона, как описано в WO 2021/180935, либо вычисляются в соответствии с множеством элементарных пространственных секторов, как описано относительно фиг. 7, при этом пространственно протяженный источник звука представляется посредством двух или

5

более элементарных пространственных секторов. Обработка для каждого аудиоблока проиллюстрирована на фиг. 11, иллюстрирующем общую блок-схему предпочтительного варианта осуществления, реализующего первый аспект, второй аспект и третий аспект вместе. Для каждого блока аудиосигналов, (изменяющиеся во времени) целевые сигнальные метки для целевой пространственной

10

области, принадлежащей SESS, определяются и применяются к двум входным сигналам в каскаде синтеза бинауральных сигнальных меток, чтобы формировать бинауральные выходные L- и R-сигналы.

Целевые бинауральные сигнальные метки вычисляются следующим образом:

Пространственные секторы, принадлежащие SESS, вычисляются с учетом положения и ориентации слушателя и SESS, а также геометрии SESS, (например, с использованием алгоритма проецирования или анализа по трассировке лучей).

15

В частности, находятся пространственные секторы, принадлежащие частям SESS, которые должны взвешиваться для моделирования таких эффектов, как загораживание и/или ослабление в зависимости от расстояния и т.д. Может быть предусмотрено несколько пространственных областей, которые требуют различных характеристик ослабления/частотного отклика; соответствующие секторы обрабатываются в каждой области отдельно, принадлежащей различным так называемым «классам секторов» (например, «незагороженный», «модификация/загораживание #1»... «модификация/загораживание номер #n»).

20

Сохраненные (ко)-вариационные члены для секторов в каждом классе секторов суммируются. Затем суммированные данные (ко)вариации секторов различных классов секторов взвешиваются согласно требуемой функции передачи для каждого класса секторов. В частности, данные (ко)вариации этого класса секторов умножаются на (частотно-зависимую) функцию передачи энергии (квадрат коэффициента амплитудного масштабирования/амплитудно-частотного отклика), принадлежащую этому классу.

25

30

Взвешенные дисперсионные члены для всех классов секторов SESS суммируются в полные (взвешенные) (ко)-вариационные члены.

Целевые сигнальные метки с использованием модифицированных/взвешенных полных (ко)-вариационных членов вычисляются с использованием уравнений (23)-(27).

35

Конечно, также данные (ко)вариации каждого сектора могут взвешиваться отдельно и затем суммироваться, вместо сначала выполнения частичного суммирования в классах секторов, взвешивания по одному разу для каждого класса секторов и конечного суммирования. Тем не менее, вышеописанный подход представляет собой предпочтительный вариант осуществления вследствие своей более высокой

40

эффективности. Преимущества вариантов осуществления изобретения по сравнению с уровнем техники обеспечивают очень эффективный и более реалистичный рендеринг размерных источников (SESS), небольшой размер таблицы поиска и/или способность включать в себя эффекты рендеринга (такие как частичное загораживание или ослабление в зависимости от расстояния), которые изменяют частотный отклик в выбранных пространственных частях размерного источника (SESS).

45

Предпочтительные примеры относятся к модулю рендеринга, который использует в качестве вводов один или более каналов передачи сигналов, геометрию, размер и

ориентацию пространственно протяженного источника звука (SESS) и набор HRTF и оснащается возможностями бинаурального рендеринга пространственно протяженных источников звука (т.е. обеспечивает два выходных сигнала).

5 Дополнительно предпочтительные модули рендеринга или устройства и способы синтеза SPESS содержат, помимо или вместо вышеуказанного, каскад вычисления целевых сигнальных меток (например, для вычисления требуемых интерауральных целевых сигнальных меток) и каскад синтеза сигнальных меток (например, для преобразования входного сигнала(ов) в подготовленные посредством бинаурального рендеринга сигналы с требуемыми целевыми сигнальными метками).

10 Дополнительно предпочтительные модули рендеринга или устройства и способы синтеза SPESS содержат, помимо или вместо вышеуказанного, использование таблицы поиска, которая содержит предварительно вычисленные данные для бинаурального рендеринга SESS и обеспечивается/предварительно вычисляется для различных полос частот в зависимости от набора HRTF.

15 Дополнительно предпочтительные модули рендеринга или устройства и способы синтеза SPESS содержат, помимо или вместо вышеуказанного, таблицу поиска, которая организуется с возможностью сохранения (ко)-вариационных членов для каждого пространственного сектора (например, l-(левой) дисперсии, r-(правой) дисперсии, ковариации lr).

20 В другом предпочтительном варианте осуществления: пространственные секторы задаются как диапазоны изменения азимута/подъема.

В других предпочтительных вариантах осуществления, размеры пространственных секторов выбираются относительно разрешения человеческих слуховых способностей к пространственной локализации (например, являются более широкими в подъемном, 25 чем в азимутальном направлении).

В других предпочтительных вариантах осуществления, вычисление целевых сигнальных меток бинаурального рендеринга выполняется на основе суммированных дисперсионных членов пространственных секторов, принадлежащих SESS.

30 В других предпочтительных вариантах осуществления, модификация рендеринга различных пространственных областей SESS (например, для моделирования на основе загоразивания или расстояния) достигается посредством использования модифицированных дисперсионных членов из таблицы поиска, а не первоначально сохраненного дисперсионного члена.

35 В других предпочтительных вариантах осуществления, модификация осуществляется посредством умножения дисперсионных членов на коэффициент ослабления энергии, принадлежащий пространственному сектору.

В других предпочтительных вариантах осуществления, этот коэффициент ослабления является частотно-зависимым (например, чтобы моделировать эффекты нижних частот вследствие частичного загоразивания).

40 Дополнительный вариант осуществления относится к потоку битов, который включает в себя следующую информацию: размер, положение и ориентация объекта, а также форма сигнала и геометрия загоразивающих объектов.

Ниже описан дополнительный предпочтительный вариант осуществления, разрабатываемый в настоящее время для MPEG I ISO 23090-4.

45 Этот вариант осуществления синтезирует один или более пространственно протяженных источников звука (SESS) для воспроизведения в наушниках для источников объектов, которые имеют ассоциированный флаг `objectSourceHasExtent`, заданный равным 1. Соответствующие параметры для источника объектов идентифицируются

посредством `objectSourceExtentId`.

Синтез основан на описании SESS посредством (в идеале) бесконечного числа декоррелированных точечных источников, распределенных по всему пространственному диапазону протяженности источника. Посредством непрерывного проецирования геометрии SESS в направлении к текущему положению слушателя, диапазон, покрытый 5 посредством упомянутой геометрии, может идентифицироваться каждый кадр и обновляться в реальном времени. Другими словами, геометрия проецируется на сферу, представляющую виртуальное пространство для прослушивания пользователя, каждый кадр. Кроме того, пространственные секции, занимаемые посредством проецируемой 10 геометрии на сфере, представляют собой секции, включенные в аурализацию SESS.

SESS задается пользователем во входном формате для кодера (EIF). С учетом требуемого диапазона протяженностей источников, SESS синтезируется с использованием двух декоррелированных входных сигналов. Эти входные сигналы обрабатываются таким образом, что перцепционно важные слуховые сигнальные метки 15 синтезируются. Это включает в себя следующие интерауральные сигнальные метки: интерауральная взаимная корреляция (IACC), интерауральные разности фаз (IAPD) и интерауральные разности уровней (IALD). Помимо этого, моноауральные спектральные сигнальные метки воспроизводятся. Это проиллюстрировано на фиг. 12.

Элементы данных и переменные

20 `itemStore` - локальный указатель на объект `RenderItemStore`

`B` - размер блока

`Fs` - частота дискретизации

`extentProcessors` - преобразование из идентификатора элемента в его экземпляр `extentProcessor`

25 `extentDownmixItem` - RI, который сохраняет конечный вывод бинаурального сигнала всей протяженности.

Описание каскада

Чтобы сокращать вычислительные затраты в реальном времени, отдельные точки HRTE назначаются в предварительно заданные таблицы-сетки, которые разделяют 30 виртуальную сферу прослушивания слушателя на равномерно распределенные области. Во время инициализации выполняется N-точечное DFT для получения N/2+1 частотных компонентов для каждой HRIR, где N является ее длиной. Затем три промежуточных значения для каждой сетки получаются посредством интегрирования данных всех точек HRTE, которые являются усилениями левых и правых каналов, внутри 35 ненормализованной IACC. Помимо этого, также сохраняется число точек данных HRTE, включенных в каждую сетку,. Они используются для вычисления конечных сигнальных меток в реальном времени.

Усиления обоих каналов для каждой сетки вычисляются с помощью уравнения 28 и 29, где $A_{l,n}$ и $A_{r,n}$ являются абсолютной величиной левой и правой HRTE, соответственно, 40 N является числом точек HRTE, которые находятся в пределах этой сетки:

$$G_l(\omega) = \sum_{n=1}^N A_{l,n}^2 \quad (28)$$

$$G_r(\omega) = \sum_{n=1}^N A_{r,n}^2 \quad (29)$$

45 Ненормализованная IACC для каждой сетки вычисляется с помощью уравнения 30, где ϕ_l и ϕ_r является фазой левой и правой HRTE, соответственно:

$$IACC(\omega) = \sum_{n=1}^N A_{l,n}^2 \cdot A_{r,n}^2 \cdot e^{j(\phi_{l,n} - \phi_{r,n})} \quad (30)$$

Процедуры в уравнениях 28-30 выполняются перед фактической обработкой заранее и соответствуют этапам 800, 810 по фиг. 8, и результаты этих обработок представляют собой данные, предпочтительно сохраненные в устройстве 2000 или 200 хранения данных, на соответствующих чертежах.

В ходе обработки в реальном времени, каждый уникальный протяженный источник звука формируется и управляется посредством процессора протяженности. Для каждого кадра, каждый активный процессор принимает буфер аудиовыборок и метаданных, указывающих, как следует синтезировать протяженный источник звука. Существуют две отдельных цепочки обработки: обработка метаданных в подпроцессе обновления и аудиообработка в подпроцессе аудиоборботки. Они описаны, соответственно, в следующих разделах, и их результаты комбинируются в конце второй цепочки для формирования бинаурального аудиовывода.

Вычисления, выполняемые в подпроцессе обновления

Для каждого уникального протяженного источника звука, одна или более несущих метаданных, в форме RI (элементов рендеринга), формируются посредством каскада обработки загораживания (например, соответствующего блоку 4000).

Этот каскад 4000 проходит в контуре по всем входящим RI и назначает релевантные метаданные протяженности соответствующему процессору. Если одна из пространственных секций из предварительно заданной таблицы покрывается и должна включаться для аурализации протяженности в этом кадре, то входящие метаданные должны содержать коэффициент усиления (элементы 4010, 4020, 4030 по фиг. 11) и список усиления, соответствующих некоторым предварительно заданным частотным элементам разрешения для него. Посредством выбора (например, 4000), взвешивания (например, 5020) и в конечном счете накопления (например, 5040) сохраненных промежуточных данных с усилением и EQ, достигается формирование произвольной формы протяженного источника звука с любой формой и степенью загораживания (размером/материалом).

Конечный фильтр получается посредством следующих этапов: После интегрирования (или накопления) всех точек сетки, указываемых в RI (элементе рендеринга), усиление левого и правого канала и IACC (например, данные дисперсии и ковариации) нормализуются с общим взвешенным числом точек данных HRTF:

$$G_l(\omega) = \frac{\sum_{n=1}^N G_{grid,l,n}}{N_{weighted}} \quad (31)$$

$$G_r(\omega) = \frac{\sum_{n=1}^N G_{grid,r,n}}{N_{weighted}} \quad (32)$$

$$IACC(\omega) = \frac{\sum_{n=1}^N IACC_{grid,n}(\omega)}{N_{weighted}} \quad (33)$$

Процедуры в уравнениях 31-33 соответствуют блоку 5040.

Частотно-зависимые H_α and H_β вычисляются с использованием нормализованной IACC:

$$H_\beta(\omega) = \sqrt{\frac{1}{2} (1 - \sqrt{1 - |IACC(\omega)|^2})} \quad (34)$$

$$H_\alpha(\omega) = \sqrt{1 - H_\beta^2(\omega)} \quad (35)$$

Вычисление в блоке 5060 соответствует обработке уравнений 34 и 35 в варианте осуществления.

Конечные стереофильтры 3210, 3220, 3230, 3240 получаются с использованием H_α и H_β , усиления левого и правого каналов (G_L и G_R) и фаза, извлеченная из точки HRTF, соответствуют центру протяженности ($phase_L$ и $phase_R$):

$$F_{1,L}(\omega) = H_\alpha(\omega) \cdot G_L(\omega) \cdot e^{j \cdot phase_L(\omega)} \quad (36)$$

$$F_{1,R}(\omega) = H_\beta(\omega) \cdot G_R(\omega) \cdot e^{j \cdot phase_R(\omega)} \quad (37)$$

$$F_{2,L}(\omega) = H_\beta(\omega) \cdot G_L(\omega) \cdot e^{j \cdot phase_L(\omega)} \quad (38)$$

$$F_{2,R}(\omega) = H_\alpha(\omega) \cdot G_R(\omega) \cdot e^{j \cdot phase_R(\omega)} \quad (39)$$

Вычисления блоков 36-39 предпочтительно также выполняются в блоке 5060.

Вычисления, выполняемые в подпроцессе аудиоборботки

Входной моносигнал сначала подается в декоррелятор 3100 для получения двух декоррелированных версий. Может использоваться декоррелятор MPEG-I или любой другой декоррелятор, например, декоррелятор, проиллюстрированный на фиг. 10.

После этого, каждый из двух декоррелированных сигналов свертывается с помощью соответствующих стереофильтров 3210, 3220, 3230, 3240, вычисленных в подпроцессе обновления, что приводит к четырем каналам вывода. Далее должно выполняться перекрестное микширование 3250, 3260, чтобы формировать конечный бинауральный вывод.

Уравнения (40) и (41) задают процесс (фильтрации и) микширования, где S_1 и S_2 означают два декоррелированных сигнала, и F_1 и F_2 представляют собой два стереофильтра (для левого и правого, соответственно), вычисленные в секции обработки метаданных. Фиг. 13 является блок-схемой последовательности сигналов для процесса. Фильтр, проиллюстрированный на фиг. 13, является аналогичным фильтру по фиг. 9.

$$S_L(\omega) = F_{1,L}(\omega) \cdot S_1(\omega) + F_{2,L}(\omega) \cdot S_2(\omega) \quad (40)$$

$$S_R(\omega) = F_{1,R}(\omega) \cdot S_1(\omega) + F_{2,R}(\omega) \cdot S_2(\omega) \quad (41)$$

Обработка в соответствии с уравнениями 40 и 41 предпочтительно выполняется в блоке 3000 аудиопроцессора или синтеза бинауральных сигнальных меток по фиг. 11 или 300 по фиг. 4, 5, 6.

Фиг. 7 иллюстрирует схематичное представление диапазона рендеринга для слушателя. Диапазон рендеринга примерно представляет собой сферу, которая центрируется вокруг пользователя. Следовательно, пользователь или слушатель (не проиллюстрирован на фиг. 7) расположен в центре сферы, и диапазон рендеринга, соответствующий этой сфере вокруг слушателя, может считаться «привязанным» к руке пользователя. Следовательно, когда пользователь изменяет своё положение в одном из горизонтального направления, вертикального направления или направления глубины (x , y , z), сфера перемещается в соответствии с перемещением пользователя относительно пространственно протяженного источника звука, который может считаться фиксированным относительно пользователя. Кроме того, когда пользователь перемещает руку, смотря вверх, смотря вниз или смотря вбок, сфера, представляющая диапазон рендеринга для слушателя, также перемещается вверх, вниз или вбок, т.е. также выполняет «перемещение», которое пользователь применяет к своей голове без перемещения в горизонтальном направлении, вертикальном направлении или направлении глубины. Таким образом, сферический диапазон рендеринга для слушателя может считаться видом «шлема», всегда придерживающимся перемещения головы

пользователя или слушателя во всех 6 степенях свободы.

Эта сфера разделяется на отдельные элементарные пространственные секторы, которые могут быть разнесены и в силу этого иметь различные размеры относительно угла азимута и подъема, чтобы отражать психоакустические выявленные сведения. В частности, диапазон рендеринга содержит сферу или часть сферы вокруг слушателя, и каждый элементарный пространственный сектор, проиллюстрированный на фиг. 7, например, имеет размер азимута и размер подъема. В частности, размер азимута и размер подъема элементарных пространственных секторов отличаются друг от друга, так что размер азимута является более точным для элементарного пространственного сектора непосредственно перед слушателем, по сравнению с размером азимута элементарного пространственного сектора больше вбок от слушателя, и/или размер азимута снижается в направлении вбок от слушателя, и/или размер подъема элементарного пространственного сектора меньше размера азимута этого сектора.

Следовательно, аспекты изобретения базируются на центрированном на пользователе представлении, которое перемещается с пользователем относительно пространственно протяженного источника звука, и голова пользователя находится в центре пространства, и сфера или часть сферы является диапазоном рендеринга.

Процессор 4000 идентификации секторов теперь определяет то, какие различные элементарные пространственные секторы представляют пространственно протяженный источник звука, проиллюстрированный на фиг. 7 в 7 000. В этом примере, например, через алгоритм трассировки лучей, начинающийся с центра этой сферы и указывающий на SESS 7000, определяется, что четыре элементарных пространственных сектора (ESS), указанные как «1», «2», «3» и «4» на фиг. 7, «принадлежат» SESS 7000 в конкретной ориентации и положении пользователя относительно SESS 7000. Следовательно, предполагается, что звуковое поле, испускаемое посредством SESS 7000, которые фактически достигает ушей пользователя, проходит через эти четыре ESS. Кроме того, загораживающий объект 7010 также проиллюстрирован на фиг. 7, и для целей примера, предполагается, что элементарный пространственный сектор (ESS 1) полностью загораживается, элементарный пространственный сектор 2 (ESS2) частично загораживается, и ESS 3, 4 не загораживаются посредством загораживающего объекта.

Следовательно, обращаясь к фиг. 11, элементарные пространственные секторы 1, 2 соответствуют элементу 4010, элементарный пространственный сектор 1 соответствует элементу 4020, и элементарный пространственный сектор 2 соответствует элементу 4030 по фиг. 11. Альтернативно, может определяться то, что частично загороженный сектор также принадлежит к тому же классу, что и класс полностью загороженного сектора, либо, если сектор загорожен только с очень небольшой частью, то также можно определить, что сектор, имеющий загораживание ниже определенного порогового значения, также определяется как вообще не загороженный.

Хотя на фиг. 7 проиллюстрировано то, что элементарные пространственные секторы и при необходимости степень загораживания для характеристики загораживания или модификации секторов являются равными для обеих ушей, т.е. для левого и правого, также может возникать такой случай, в котором число и/или идентификация элементарных пространственных секторов отличаются для левого и для правого уха. Это может легко иметь место, когда SESS находится достаточно близко к пользователю, и SESS расположен больше в середине между обоими ушами, а не на одной или на другой стороне.

Кроме того, для определения проекции SESS на диапазон рендеринга для слушателя, т.е. для примерной сферы, могут выполняться процедуры, отличные от алгоритмов

трассировки лучей. Кроме того, SESS 7000 не должен обязательно быть фиксированным. SESS также может быть динамическим, т.е. может перемещаться со временем. Затем положение SESS относительно пользователя должно быть задано, и, после этого, для определенной точки во времени/для определенного кадра волновой формы сигнала SESS, определяются соответствующие элементарные пространственные секторы для 5 левой стороны и правой стороны слушателя для фактического положения головы слушателя, и далее сигнальные метки вычисляются так, как проиллюстрировано относительно блоков 5020-5060 на фиг. 11.

Кроме того, здесь следует отметить, что диапазон рендеринга не должен обязательно 10 представлять собой полную сферу. Он может только содержать часть сферы. Кроме того, диапазон рендеринга не должен обязательно быть сферическим. Он также может быть цилиндрическим, либо он также может иметь форму многоугольника при условии, что он покрывает определенную трехмерную часть пространства вокруг слушателя.

Относительно размеров элементарных пространственных секторов, следует 15 подчеркнуть, что элементарные пространственные секторы могут быть довольно небольшими, так что, для определения сохраненных элементов данных рендеринга, только одна HRTF, указываемая с амплитудой и фазой вместо суммирования по определенному числу (например, как проиллюстрировано в уравнении 20, уравнении 21 и уравнении 22 или в уравнении 28-30) является достаточной. Тем не менее, когда 20 используются элементарные пространственные секторы, которые имеют определенную размерность таким образом, что размер устройства хранения данных, сохраняющего элементы данных рендеринга для каждого элементарного пространственного сектора, уменьшается, определение элементов данных рендеринга, сохраненных в устройстве хранения данных для каждого элементарного пространственного сектора, может 25 выполняться в соответствии с уравнениями 20-22 или 28-30, при этом только HRTF, принадлежащие конкретному элементарному пространственному сектору, суммируются для получения фактических данных (ко)вариации для определенной частоты и для этого элементарного пространственного сектора.

Следует отметить, что конкретное преимущество этой процедуры состоит в том, что 30 все эти вычисления не должны обязательно выполняться во время выполнения. Вместо этого, как только определенное разделение диапазона рендеринга на определенную сетку элементарных пространственных секторов или точек сетки определяется, затем сохраненные данные для каждого отдельного или элементарного пространственного сектора могут вычисляться и сохраняться, и для определенной инициализации с 35 определенной сеткой, единственная процедура, осуществляемая во время выполнения, заключается в том, чтобы загружать соответствующие предварительно вычисленные данные для этой сетки в устройство хранения данных или таблицу поиска.

Единственная процедура, которая обязательно должна выполняться во время 40 выполнения, представляет собой идентификацию элементарных пространственных секторов, принадлежащих пространственно протяженному источнику звука для конкретной ориентации/положения пользователя, и потенциально необходимое взвешивание вследствие загромождающих объектов и, после этого, конечное полное суммирование, соответствующее блоку 5040 на фиг. 11, которое затем не стоит ничего для конечного вычисления целевых сигнальных меток в блоке 5060. Следовательно, 45 необходимые операции вычисления во время выполнения являются очень ограниченными и являются очень небольшими по сравнению с операциями вычисления, требуемыми для определения элементов данных рендеринга для элементарных пространственных секторов, т.е. для определенной сетки.

Кроме того, следует отметить, что устройство хранения данных для определенной сетки не зависит от положения/ориентации пользователя, поскольку, в случае изменения положения или характеристики SESS или в случае изменения ориентации/положения пользователя, изменяются только идентифицированные элементарные пространственные секторы, но не данные, сохраненные для элементарных пространственных секторов, которые представляют сетку. Другими словами, изменяются только идентификационные номера для элементарных пространственных секторов, но не данные для элементарного пространственного сектора, имеющего определенный идентификационный номер.

Ниже описана фиг. 8, которая иллюстрирует предпочтительную процедуру для одного или нескольких аспектов изобретения.

На этапе 800 определяется или инициализируется диапазон рендеринга, такой как сфера. Результат, например, представляет собой сферу с определенными точками сетки или элементарными пространственными секторами. В блоке 810, элементы данных рендеринга, например, данные (ко)вариации сохраняются в устройстве хранения данных, таком как таблица поиска для всех элементарных пространственных секторов в диапазоне рендеринга.

Далее, на этапе 820, выполняется идентификация секторов, осуществляемая посредством блока 4000. Следовательно, один или более элементарных пространственных секторов, принадлежащих пространственно протяженному источнику звука, определяются на основе данных SESS и данных положения/ориентации слушателя, вводимых в блок 820. Результат блока 820 представляет собой один или более элементарных пространственных секторов.

В блоке 830, суммирование элементов данных рендеринга для множества элементарных пространственных секторов, например, с или без взвешивания выполняется, как проиллюстрировано посредством блока 5040.

В блоке 840, вычисляются целевые данные рендеринга, такие как IACC, IALD, IAPD, GL, GR, что выполняется посредством блока 5060.

В блоке 850, целевые данные рендеринга применяются к аудиосигналу пространственно протяженного источника звука, как проиллюстрировано, например, также посредством блока 3000 аудиопроцессора или блока 3000 синтеза бинауральных сигнальных меток по фиг. 11.

В соответствии с первым аспектом настоящего изобретения, сфера рендеринга реализуется так, как проиллюстрировано на фиг. 7, т.е. элементарные пространственные секторы, покрывающие диапазон рендеринга для слушателя, определяются, и процессор идентификации секторов задает набор элементарных пространственных секторов, например, два или более элементарных пространственных сектора для пространственно протяженного источника звука. Тем не менее, только в предпочтительном варианте осуществления сохраненные элементы данных рендеринга представляют собой данные дисперсии или ковариации. Вместо этого, другие элементы данных, необходимые для рендеринга, также могут сохраняться и комбинироваться посредством модуля вычисления целевых данных. Кроме того, эта процедура также не требует обязательно обработки модификации, но предпочтительно выполняет обработку модификации.

В соответствии со вторым аспектом настоящего изобретения, требуется определение потенциально модифицирующего объекта и определение ограниченного модифицированного пространственного сектора на основе идентификации потенциально модифицирующих объектов. Тем не менее, для этой процедуры, диапазон рендеринга не должен обязательно иметь такие размеры, как проиллюстрировано на фиг. 7, т.е. с отдельными элементарными пространственными секторами, имеющими отдельные

сохраненные элементы данных. Вместо этого, диапазон рендеринга также может реализовываться так, как проиллюстрировано в других реализациях, например, в реализации, проиллюстрированной в WO 2021/180935. Кроме того, для определения и для учета объектов для модификации, не обязательно имеет место то, что сохраненные
5 элементы данных рендеринга представляют собой данные дисперсии/ковариации. Вместо этого, также могут использоваться другие данные рендеринга, например, проиллюстрированные в качестве сохраненных данных в WO 2021/180935.

Относительно третьего аспекта, определение диапазона рендеринга, как проиллюстрировано на фиг. 7, не обязательно требуется. Вместо этого, другое
10 определение, например, определения диапазона рендеринга, как проиллюстрировано в WO 2021/180935, может использоваться для одного или более ограниченных пространственных секторов. Тем не менее, ограниченный пространственный сектор предпочтительно реализуется в качестве элементарного пространственного сектора, показанного на фиг. 7. Кроме того, например, для целей использования данных
15 дисперсии/ковариации в качестве сохраненных данных, конкретная обработка с модифицирующими/загораживающими объектами также не представляет собой требуемый признак, но является предпочтительной, как пояснено выше относительно блока 830 на фиг. 8.

Далее обобщаются дополнительные варианты осуществления, относящиеся к первому
20 аспекту.

Варианты осуществления относятся к устройству для синтезирования пространственно протяженного источника звука (SESS), содержащему: устройство хранения данных для сохранения элементов данных рендеринга для различных элементарных пространственных секторов, покрывающих диапазон рендеринга для
25 слушателя; процессор идентификации секторов для идентификации, из различных элементарных пространственных секторов, набора элементарных пространственных секторов, принадлежащих пространственно протяженному источнику звука, на основе данных слушателя и данных пространственно протяженных источников звука; модуль вычисления целевых данных для вычисления целевых данных рендеринга из элементов
30 данных рендеринга для набора элементарных пространственных секторов; и аудиопроцессор для обработки аудиосигнала, представляющего пространственно протяженный источник звука, с использованием целевых данных рендеринга.

В дополнительных вариантах осуществления, устройство хранения данных выполнено с возможностью сохранения в качестве элементов данных рендеринга для каждого
35 элементарного пространственного сектора по меньшей мере одного из левого элемента данных дисперсии, связанного с левыми данными передаточной функции восприятия звука человеком, правого элемента данных дисперсии, связанного с правыми данными передаточной функции восприятия звука человеком (HRTF), и элемента данных ковариации, связанного с левыми данными HRTF и правыми данными HRTF, при этом
40 модуль вычисления целевых данных выполнен с возможностью суммирования левых элементов данных дисперсии для набора элементарных пространственных секторов или правых элементов данных дисперсии для набора элементарных пространственных секторов, или элементов данных ковариации для набора элементарных пространственных секторов, соответственно, для получения по меньшей мере одного суммированного элемента, при этом модуль вычисления целевых данных выполнен с
45 возможностью вычисления по меньшей мере одной сигнальной метки рендеринга в качестве целевых данных рендеринга по меньшей мере из одного суммированного элемента, и при этом аудиопроцессор выполнен с возможностью обработки

аудиосигнала с использованием по меньшей мере одной сигнальной метки рендеринга.

В дополнительных вариантах осуществления, процессор идентификации секторов выполнен с возможностью применения алгоритма проецирования или анализ по трассировке лучей для определения набора элементарных пространственных секторов или использования в качестве данных слушателя положения слушателя или ориентации слушателя либо использования в качестве данных пространственно протяженных источников звука (SESS) ориентации SESS, положения SESS или информации относительно геометрии SESS.

В дополнительных вариантах осуществления, процессор идентификации секторов выполнен с возможностью приёма из описания аудиосцены информации загораживания относительно потенциально загораживающего объекта и определения на основе информации загораживания конкретного пространственного сектора из набора элементарных пространственных секторов в качестве загораживающего сектора, и при этом модуль вычисления целевых данных выполнен с возможностью применения функции загораживания к элементам данных рендеринга, сохраненным для загораживающего сектора, таким образом, чтобы получить модифицированные данные и использовать модифицированные данные для вычисления целевых данных рендеринга.

В дополнительных вариантах осуществления, функция загораживания представляет собой функцию нижних частот, имеющую различные значения ослабления для различных частот, и при этом элементы данных рендеринга представляют собой элементы данных для различных частот, и при этом модуль вычисления целевых данных выполнен с возможностью взвешивания для нескольких частот элемента данных для определенной частоты со значением ослабления для определенной частоты для получения модифицированных данных рендеринга.

В дополнительных вариантах осуществления, процессор идентификации секторов выполнен с возможностью определения того, что другой элементарный пространственный сектор из набора элементарных пространственных секторов, определенных для загораживающего объекта, не загораживается посредством потенциального загораживающего объекта, и при этом модуль вычисления целевых данных выполнен с возможностью комбинирования модифицированных данных из загораживающего сектора и элементов данных рендеринга другого сектора без модификации с использованием функции загораживания либо с модификацией посредством другой функции модификации для получения целевых данных рендеринга.

В дополнительных вариантах осуществления процессор идентификации секторов выполнен с возможностью определения первого элементарного пространственного сектора из набора элементарных пространственных секторов как имеющего первую характеристику и определения второго элементарного пространственного сектора из набора элементарных пространственных секторов как имеющего вторую отличающуюся характеристику, и при этом модуль вычисления целевых данных выполнен с возможностью неприменения функции модификации к первому элементарному пространственному сектору и применения функции модификации к второму элементарному пространственному сектору либо применения первой функции модификации к первому элементарному пространственному сектору и применения второй функции модификации ко второму элементарному пространственному сектору, причем вторая функция модификации отличается от первой функции модификации.

В дополнительных вариантах осуществления, первая функция модификации является частотно-избирательной, и вторая функция модификации является постоянной по частоте, либо при этом первая функция модификации имеет первую частотно-

избирательную характеристику, и при этом вторая функция модификации имеет вторую частотно-избирательную характеристику, отличающуюся от первой частотно-избирательной характеристики, либо при этом первая функция модификации имеет первую характеристику ослабления, и вторая функция модификации имеет вторую отличающуюся характеристику ослабления, и при этом модуль вычисления целевых данных выполнен с возможностью выбора или регулировки функции модификации из первой функции модификации и второй функции модификации на основе расстояния между первым элементарным пространственным сектором или вторым элементарным пространственным сектором и слушателем либо на основе характеристики объекта, расположенного между слушателем и соответствующим элементарным пространственным сектором.

В дополнительных вариантах осуществления, процессор идентификации секторов выполнен с возможностью классификации набора элементарных пространственных секторов на различные классы секторов на основе характеристик, ассоциированных с элементарными пространственными секторами, при этом модуль вычисления целевых данных выполнен с возможностью комбинирования элементов данных рендеринга элементарных пространственных секторов в каждом классе для получения комбинированного результата для каждого класса, если более одного элементарного пространственного сектора находится в классе, и применения конкретной функции модификации, ассоциированной по меньшей мере с одним классом, к комбинированному результату этого класса для получения модифицированного комбинированного результата для этого класса, или применения конкретной функции модификации, ассоциированной по меньшей мере с одним классом, к одному или более элементам данных одного или более элементарных пространственных секторов каждого класса для получения модифицированных элементов данных, и комбинирования модифицированных элементов данных элементарных пространственных секторов в каждом классе для получения модифицированного комбинированного результата для этого класса, комбинирования комбинированного результата, либо, при наличии, модифицированного комбинированного результата для каждого класса для получения полного комбинированного результата, и использования полного комбинированного результата в качестве целевых данных рендеринга или вычисления целевых данных рендеринга из полного комбинированного результата.

В дополнительных вариантах осуществления, характеристика для элементарного пространственного сектора определяется как представляющий собой одно из группы, содержащей загороженный элементарный пространственный сектор, включающий в себя первую характеристику загораживания, загороженный элементарный пространственный сектор, включающий в себя вторую характеристику загораживания, отличающуюся от первой характеристики загораживания, незагороженный элементарный пространственный сектор, имеющий первое расстояние до слушателя, и незагороженный элементарный пространственный сектор, имеющий второе расстояние до слушателя, при этом второе расстояние отличается от первого расстояния.

В дополнительных вариантах осуществления, модуль вычисления целевых данных выполнен с возможностью модификации или комбинирования частотно-зависимых параметров дисперсии или ковариации в качестве элементов данных рендеринга для получения в качестве полного комбинированного результата полной комбинированной дисперсии или полного комбинированного параметра ковариации, и вычисления по меньшей мере одного из сигнальной метки интерауральной когерентности, сигнальной метки интерауральной разности уровней, сигнальной метки интерауральной разности

фаз, первого бокового усиления или второго бокового усиления в качестве целевых данных рендеринга.

В дополнительных вариантах осуществления, аудиопроцессор выполнен с возможностью выполнения по меньшей мере одного из регулирования межканальной когерентности, регулирования межканальной разности фаз, регулирования межканальной разности уровней с использованием соответствующих сигнальных меток в качестве целевых данных рендеринга.

В дополнительных вариантах осуществления, диапазон рендеринга содержит сферу или часть сферы вокруг слушателя, при этом диапазон рендеринга привязывается к положению слушателя или к ориентации слушателя, и при этом каждый элементарный пространственный сектор имеет размер азимута и размер подъема.

В дополнительных вариантах осуществления, размер азимута и размер подъема элементарных пространственных секторов отличаются друг от друга, так что размер азимута является более точным для элементарного пространственного сектора непосредственно перед слушателем по сравнению с размером азимута элементарного пространственного сектора больше вбок от слушателя, либо при этом размер азимута снижается в направлении вбок от слушателя, либо при этом размер подъема элементарного пространственного сектора меньше размера азимута этого сектора.

Далее обобщаются дополнительные варианты осуществления, относящиеся ко второму аспекту.

Вариант осуществления для устройства для синтезирования пространственно протяженного источника звука содержит: входной интерфейс для приема описания аудиосцены, причем описание аудиосцены содержит данные пространственно протяженных источников звука по пространственно протяженному источнику звука и данные для модификации по потенциально модифицирующему объекту, и для приема данных слушателя; процессор идентификации секторов для идентификации ограниченного модифицированного пространственного сектора для пространственно протяженного источника звука в диапазоне рендеринга для слушателя, причем диапазон рендеринга для слушателя больше ограниченного модифицированного пространственного сектора, на основе данных пространственно протяженных источников звука и данных слушателя, и данных для модификации; модуль вычисления целевых данных для вычисления целевых данных рендеринга из одного или более элементов данных рендеринга, принадлежащих модифицированному ограниченному пространственному сектору; и аудиопроцессор для обработки аудиосигнала, представляющего пространственно протяженный источник звука, с использованием целевых данных рендеринга.

В дополнительных вариантах осуществления, данные для модификации представляют собой данные затемнения, и при этом потенциально модифицирующий объект представляет собой потенциально загораживающий объект.

В дополнительных вариантах осуществления, потенциально модифицирующий объект имеет ассоциированную функцию модификации, при этом один или более элементов данных рендеринга являются частотно-зависимыми, при этом функция модификации является частотно-избирательной, и при этом модуль вычисления целевых данных выполнен с возможностью применения частотно-избирательной функции модификации к одному или более частотно-зависимым элементам данных рендеринга.

В дополнительных вариантах осуществления частотно-избирательная функция модификации имеет различные значения для различных частот, и при этом частотно-зависимые один или более элементов данных рендеринга имеют различные значения

для различных частот, и при этом модуль вычисления целевых данных выполнен с возможностью применения или умножения или комбинирования значения частотно-избирательной функции модификации для определенной частоты относительно значения одного или более элементов данных рендеринга для определенной частоты.

5 В дополнительных вариантах осуществления предложено устройство хранения данных для сохранения одного или более элементов данных рендеринга для определенного числа различных ограниченных пространственных секторов, при этом определенное число различных ограниченных пространственных секторов вместе формируют диапазон рендеринга для слушателя.

10 В дополнительных вариантах осуществления функция модификации представляет собой частотно-избирательную функцию нижних частот, и при этом модуль вычисления целевых данных выполнен с возможностью применения функции нижних частот таким образом, что значение одного или более элементов данных рендеринга на более высокой частоте ослабляется сильнее значения одного или более элементов данных рендеринга на более низкой частоте.

15 В дополнительных вариантах осуществления, процессор идентификации секторов выполнен с возможностью определения ограниченного пространственного сектора для пространственно протяженного источника звука на основе данных слушателя и данных пространственно протяженных источников звука, определения, подвергается ли по меньшей мере часть ограниченного пространственного сектора модификации посредством модифицирующего объекта, и определения ограниченного пространственного сектора в качестве модифицированного пространственного сектора, когда часть больше порогового значения, либо когда целый ограниченный пространственный сектор подвергается модификации посредством модифицирующего

25 объекта.

В дополнительных вариантах осуществления, процессор идентификации секторов выполнен с возможностью применения алгоритма проецирования или анализа по трассировке лучей для определения ограниченного пространственного сектора или использования в качестве данных слушателя положения слушателя или ориентации слушателя либо использования в качестве данных пространственно протяженных источников звука (SESS) ориентацию SESS, положения SESS или информации относительно геометрии SESS.

30

В дополнительных вариантах осуществления, диапазон рендеринга содержит сферу или часть сферы вокруг слушателя, при этом диапазон рендеринга привязывается к положению слушателя или к ориентации слушателя, и при этом модифицированный ограниченный пространственный сектор имеет размер азимута и размер подъема.

35

В дополнительных вариантах осуществления, размер азимута и размер подъема модифицированного ограниченного пространственного сектора отличаются друг от друга, так что размер азимута является более точным для модифицированного ограниченного пространственного сектора непосредственно перед слушателем по сравнению с размером азимута модифицированного ограниченного пространственного сектора больше вбок от слушателя, либо при этом размер азимута снижается в направлении вбок от слушателя, либо при этом размер подъема модифицированного ограниченного пространственного сектора меньше размера азимута модифицированного ограниченного пространственного сектора.

40

45

В дополнительных вариантах осуществления, в качестве одного или более элементов данных рендеринга, для модифицированного ограниченного пространственного сектора, используется по меньшей мере одно из левого элемента данных дисперсии, связанного

с левыми данными передаточной функции восприятия звука человеком, правого элемента данных дисперсии, связанного с правыми данными передаточной функции восприятия звука человеком (HRTF), и элемента данных ковариации, связанного с левыми данными HRTF и правыми данными HRTF.

5 В дополнительных вариантах осуществления, процессор идентификации секторов выполнен с возможностью определения набора элементарных пространственных секторов, принадлежащих пространственно протяженному источнику звука, и определения из набора элементарных пространственных секторов одного или более элементарных пространственных секторов в качестве ограниченного
10 модифицированного пространственного сектора, и при этом модуль вычисления целевых данных выполнен с возможностью модификации одного или более элементов данных рендеринга, ассоциированных с ограниченным модифицированным пространственным сектором, с использованием данных для модификации для получения комбинированных данных, и комбинирования комбинированных данных с элементами данных рендеринга
15 одного или более элементарных пространственных секторов из набора элементарных пространственных секторов, отличающихся от ограниченного модифицированного пространственного сектора и не модифицируемых или модифицируемых другим способом по сравнению с модификацией для ограниченного модифицированного пространственного сектора.

20 В дополнительных вариантах осуществления, процессор идентификации секторов выполнен с возможностью классификации набора элементарных пространственных секторов на различные классы секторов на основе характеристик, ассоциированных с элементарными пространственными секторами, при этом модуль вычисления целевых данных выполнен с возможностью комбинирования элементов данных рендеринга
25 элементарных пространственных секторов в каждом классе для получения комбинированного результата для каждого класса, если в классе находится более одного элементарного пространственного сектора, и применения конкретной функции модификации, ассоциированной по меньшей мере с одним классом, к комбинированному результату этого класса для получения модифицированного комбинированного
30 результата для этого класса, или применения конкретной функции модификации, ассоциированной по меньшей мере с одним классом, к одному или более элементам данных одного или более элементарных пространственных секторов каждого класса для получения модифицированных элементов данных, и комбинирования модифицированных элементов данных элементарных пространственных секторов в
35 каждом классе для получения модифицированного комбинированного результата для этого класса, комбинирования комбинированного результата, либо, при наличии, модифицированного комбинированного результата для каждого класса для получения полного комбинированного результата, и использования полного комбинированного результата в качестве целевых данных рендеринга или вычисления целевых данных
40 рендеринга из полного комбинированного результата.

В дополнительных вариантах осуществления, характеристика для элементарного пространственного сектора определяется как одно из группы, содержащей загороженный элементарный пространственный сектор, включающий в себя первую характеристику загораживания, загороженный элементарный пространственный сектор,
45 включающий в себя вторую характеристику загораживания, отличающуюся от первой характеристики загораживания, незагороженный элементарный пространственный сектор, имеющий первое расстояние до слушателя, и незагороженный элементарный пространственный сектор, имеющий второе расстояние до слушателя, при этом второе

расстояние отличается от первого расстояния.

В дополнительных вариантах осуществления, модуль вычисления целевых данных выполнен с возможностью модификации или комбинирования частотно-зависимых параметров дисперсии или ковариации в качестве элементов данных рендеринга для получения в качестве полного комбинированного результата полной комбинированной дисперсии или полного комбинированного параметра ковариации, и вычисления по меньшей мере одного из сигнальной метки интерауральной или межканальной когерентности, сигнальной метки интерауральной или межканальной разности уровней, сигнальной метки интерауральной или межканальной разности фаз, первого бокового усиления или второго бокового усиления в качестве целевых данных рендеринга, и при этом аудиопроцессор выполнен с возможностью обработки аудиосигнала с использованием по меньшей мере одного из сигнальной метки интерауральной или межканальной когерентности, сигнальной метки интерауральной или межканальной разности уровней, сигнальной метки интерауральной или межканальной разности фаз, первого бокового усиления или второго бокового усиления в качестве целевых данных рендеринга.

Дополнительные варианты осуществления содержат генератор аудиосцен для формирования описания аудиосцен, содержащий: генератор данных пространственно протяженных источников звука (SESS) для формирования данных SESS пространственно протяженного источника звука; генератор данных для модификации для формирования данных для модификации по потенциально модифицирующему объекту; и выходной интерфейс для формирования описания аудиосцен, содержащего данные SESS и данные для модификации.

В дополнительных вариантах осуществления, данные для модификации содержат описание функции нижних частот или геометрических данных относительно потенциально модифицирующего объекта, при этом функция нижних частот содержит значение ослабления для более высокой частоты, причем значение ослабления для более высокой частоты, представляющее значение ослабления, является более сильным по сравнению со значением ослабления для более низкой частоты, и при этом выходной интерфейс выполнен с возможностью ввода описания функции ослабления или геометрических данных относительно потенциально модифицирующего объекта в качестве данных для модификации в описание аудиосцен.

В дополнительных вариантах осуществления, генератор -данных SESS выполнен с возможностью формирования в качестве данных SESS местоположения SESS и информации относительно геометрии SESS, и при этом выходной интерфейс выполнен с возможностью ввода в качестве -данных SESS информации относительно местоположения SESS и информации относительно геометрии SESS.

В дополнительных вариантах осуществления, генератор данных SESS выполнен с возможностью формирования в качестве данных SESS информации относительно размера, относительно положения или ориентации пространственно протяженного источника звука либо данных формы сигнала для одного или более аудиосигналов, ассоциированных с пространственно протяженным источником звука, либо при этом модуль вычисления данных для модификации выполнен с возможностью вычисления в качестве данных для модификации геометрии потенциально модифицирующего объекта, например, потенциально загораживающего объекта.

Дополнительные варианты осуществления содержат описание аудиосцен, содержащее: данные пространственно протяженных источников звука и данные для модификации по одному или более потенциально модифицирующим объектам.

В дополнительных вариантах осуществления, описание аудиосцен реализуется как передаваемый или сохраненный поток битов, при этом данные пространственно протяженных источников звука представляют первый элемент потока битов, и при этом данные для модификации представляют второй элемент потока битов.

5 Далее обобщаются дополнительные варианты осуществления, относящиеся к третьему аспекту.

Вариант осуществления содержит устройство для синтезирования пространственно протяженного источника звука (SESS), содержащее: устройство хранения данных для сохранения одного или более элементов данных рендеринга для различных
10 ограниченных пространственных секторов, при этом различные ограниченные пространственные секторы расположены в диапазоне рендеринга для слушателя, при этом один или более элементов данных рендеринга для ограниченного пространственного сектора содержат по меньшей мере один из левого элемента данных дисперсии, связанного с левыми данными передаточной функции восприятия звука
15 человеком, правого элемента данных дисперсии, связанного с правыми данными передаточной функции восприятия звука человеком, и элемента данных ковариации, связанного с левыми данными передаточной функции восприятия звука человеком и правыми данными передаточной функции восприятия звука человеком; процессор идентификации секторов для идентификации одного или более ограниченных
20 пространственных секторов для пространственно протяженного источника звука в диапазоне рендеринга для слушателя на основе данных пространственно протяженных источников звука; модуль вычисления целевых данных для вычисления целевых данных рендеринга из сохраненных левых данных дисперсии, сохраненных правых данных дисперсии или сохраненных данных ковариации; и аудиопроцессор для обработки
25 аудиосигнала, представляющего пространственно протяженный источник звука, с использованием целевых данных рендеринга.

В дополнительных вариантах осуществления, устройство хранения данных выполнено с возможностью сохранения элементов данных дисперсии или элемента данных ковариации, связанных с данными передаточной функции восприятия звука человеком
30 или с данными бинаурального импульсного отклика в помещении, или с данными бинауральной передаточной функции в помещении, или с данными импульсного отклика восприятия звука человеком.

В дополнительных вариантах осуществления, один или более элементов данных рендеринга содержат значения элементов данных дисперсии или ковариации для
35 различных частот.

В дополнительных вариантах осуществления, устройство хранения данных выполнено с возможностью сохранения для каждого ограниченного пространственного сектора частотно-зависимого представления левого элемента данных дисперсии, частотно-зависимого представления правого элемента данных дисперсии и частотно-зависимого
40 представления элемента данных ковариации.

В дополнительных вариантах осуществления, модуль вычисления целевых данных выполнен с возможностью вычисления в качестве целевых данных рендеринга по
меньшей мере одного из сигнальной метки интерауральной или межканальной когерентности, сигнальной метки интерауральной или межканальной разности уровней,
45 сигнальной метки интерауральной или межканальной разности фаз, первого бокового усиления и второго бокового усиления в качестве целевых данных рендеринга, и при этом аудиопроцессор выполнен с возможностью выполнения по меньшей мере одного из регулирования межканальной или интерауральной когерентности, регулирования

интерауральной или межканальной разности фаз или регулирования интерауральной или межканальной разности уровней с использованием соответствующих сигнальных меток в качестве целевых данных рендеринга.

5 В дополнительных вариантах осуществления, модуль вычисления целевых данных выполнен с возможностью вычисления сигнальной метки интерауральной или межканальной когерентности на основе левого элемента данных дисперсии, правого элемента данных дисперсии и элемента данных ковариации либо вычисления сигнальной метки межканальной или интерауральной разности фаз на основе левого элемента данных дисперсии и правого элемента данных дисперсии, либо вычисления сигнальной метки межканальной или интерауральной разности фаз на основе элемента данных ковариации, либо вычисления левого или правого бокового усиления с использованием левого или правого элемента данных дисперсии и информации, связанной с мощностью сигналов аудиосигнала.

15 В дополнительных вариантах осуществления, модуль вычисления целевых данных выполнен с возможностью вычисления сигнальной метки интерауральной или межканальной когерентности таким образом, что значение сигнальной метки интерауральной или межканальной когерентности находится в пределах диапазона в $\pm 20\%$ от значения, полученного посредством уравнения для сигнальной метки интерауральной или межканальной когерентности, поясненной в описании изобретения, либо при этом модуль вычисления целевых данных выполнен с возможностью вычисления сигнальной метки интерауральной или межканальной разности уровней таким образом, что значение сигнальной метки интерауральной или межканальной разности уровней находится в пределах диапазона в $\pm 20\%$ от значения, полученного посредством уравнения для сигнальной метки интерауральной или межканальной разности уровней, поясненной в описании изобретения, либо при этом модуль вычисления целевых данных выполнен с возможностью вычисления сигнальной метки интерауральной или межканальной разности фаз таким образом, что значение сигнальной метки интерауральной или межканальной разности фаз находится в пределах диапазона в $\pm 20\%$ от значения, полученного посредством уравнения для сигнальной метки интерауральной или межканальной разности фаз, поясненной в описании изобретения, либо при этом модуль вычисления целевых данных выполнен с возможностью вычисления первого или второго бокового усиления таким образом, что значение первого или второго бокового усиления находится в пределах диапазона в $\pm 20\%$ от значения, полученного посредством уравнения для левого или правого бокового усиления, поясненного в описании изобретения.

В дополнительных вариантах осуществления, процессор идентификации секторов выполнен с возможностью применения алгоритма проецирования или анализа по трассировке лучей для определения одного или более ограниченных пространственных секторов в качестве набора элементарных пространственных секторов или использования в качестве данных слушателя положения слушателя или ориентации слушателя либо использования в качестве данных пространственно протяженных источников звука (SESS) ориентации SESS, положения SESS или информации относительно геометрии SESS.

45 В дополнительных вариантах осуществления, диапазон рендеринга содержит сферу или часть сферы вокруг слушателя, при этом диапазон рендеринга привязывается к положению слушателя или к ориентации слушателя, и при этом один или более ограниченных пространственных секторов имеют размер азимута и размер подъема.

В дополнительных вариантах осуществления, размер азимута и размер подъема

различных ограниченных пространственных секторов отличаются друг от друга, так что размер азимута является более точным для ограниченного пространственного сектора непосредственно перед слушателем по сравнению с размером азимута ограниченного пространственного сектора больше вбок от слушателя, либо при этом
5 размер азимута снижается в направлении вбок от слушателя, либо при этом размер подъема ограниченного пространственного сектора меньше размера азимута этого сектора.

В дополнительных вариантах осуществления, процессор идентификации секторов выполнен с возможностью определения набора элементарных пространственных
10 секторов в качестве одного или более ограниченных пространственных секторов, при этом для каждого элементарного пространственного сектора сохраняется по меньшей мере один из левого элемента данных дисперсии, правого элемента данных дисперсии и элемента данных ковариации.

В дополнительных вариантах осуществления, процессор идентификации секторов
15 выполнен с возможностью приёма из описания аудиосцены информации загораживания относительно потенциально загораживающего объекта и определения на основе информации загораживания конкретного пространственного сектора из набора элементарных пространственных секторов в качестве загораживающего сектора, и при этом модуль вычисления целевых данных выполнен с возможностью применения
20 функции загораживания к элементам данных рендеринга, сохраненным для загораживающего сектора таким образом, чтобы получить модифицированные данные и использовать модифицированные данные для вычисления целевых данных рендеринга.

В дополнительных вариантах осуществления, функция загораживания представляет собой функцию нижних частот, имеющую различные значения ослабления для различных
25 частот, и при этом элементы данных рендеринга представляют собой элементы данных для различных частот, и при этом модуль вычисления целевых данных выполнен с возможностью взвешивания для нескольких частот элемента данных для определенной частоты со значением ослабления для определенной частоты для получения модифицированных данных рендеринга.

В дополнительных вариантах осуществления, процессор идентификации секторов
30 выполнен с возможностью определения того, что другой элементарный пространственный сектор из набора элементарных пространственных секторов, определенных для загораживающего объекта, не загораживается посредством потенциального загораживающего объекта, и при этом модуль вычисления целевых
35 данных выполнен с возможностью комбинирования модифицированных данных из загораживающего сектора и элементов данных рендеринга другого сектора без модификации с использованием функции загораживания либо с модификацией посредством другой функции модификации для получения целевых данных рендеринга.

В дополнительных вариантах осуществления, процессор идентификации секторов
40 выполнен с возможностью определения первого элементарного пространственного сектора из набора элементарных пространственных секторов как имеющего первую характеристику и определения второго элементарного пространственного сектора из набора элементарных пространственных секторов как имеющего вторую отличающуюся
45 характеристику, и при этом модуль вычисления целевых данных выполнен с возможностью неприменения функции модификации к первому элементарному пространственному сектору и применения функции модификации ко второму элементарному пространственному сектору либо применения первой функции модификации к первому элементарному пространственному сектору и применения

вторую функцию модификации ко второму элементарному пространственному сектору, причем вторая функция модификации отличается от первой функции модификации.

В дополнительных вариантах осуществления, первая функция модификации является частотно-избирательной, и вторая функция модификации является постоянной по частоте, либо при этом первая функция модификации имеет первую частотно-избирательную характеристику, и при этом вторая функция модификации имеет вторую частотно-избирательную характеристику, отличающуюся от первой частотно-избирательной характеристики, либо при этом первая функция модификации имеет первую характеристику ослабления, и вторая функция модификации имеет вторую отличающуюся характеристику ослабления, и при этом модуль вычисления целевых данных выполнен с возможностью выбора или регулировки функции модификации из первой функции модификации и второй функции модификации на основе расстояния между первым элементарным пространственным сектором или вторым элементарным пространственным сектором и слушателем либо на основе характеристики объекта, расположенного между слушателем и соответствующим элементарным пространственным сектором.

В дополнительных вариантах осуществления, процессор идентификации секторов выполнен с возможностью классификации набора элементарных пространственных секторов на различные классы секторов на основе характеристик, ассоциированных с элементарными пространственными секторами, при этом модуль вычисления целевых данных выполнен с возможностью комбинирования элементов данных рендеринга элементарных пространственных секторов в каждом классе для получения комбинированного результата для каждого класса, если более одного элементарного пространственного сектора находится в классе, и применения конкретной функции модификации, ассоциированной по меньшей мере с одним классом, к комбинированному результату этого класса, для получения модифицированного комбинированного результата для этого класса, или применения конкретной функции модификации, ассоциированной по меньшей мере с одним классом, к одному или более элементам данных одного или более элементарных пространственных секторов каждого класса для получения модифицированных элементов данных, и комбинирования модифицированных элементов данных элементарных пространственных секторов в каждом классе для получения модифицированного комбинированного результата для этого класса, комбинирования комбинированного результата, либо, при наличии, модифицированного комбинированного результата для каждого класса, для получения полного комбинированного результата, и использования полного комбинированного результата в качестве целевых данных рендеринга или вычисления целевых данных рендеринга из полного комбинированного результата.

В дополнительных вариантах осуществления, характеристика для элементарного пространственного сектора определяется как представляющий собой одно из группы, содержащей загороженный элементарный пространственный сектор, включающий в себя первую характеристику загороживания, загороженный элементарный пространственный сектор, включающий в себя вторую характеристику загороживания, отличающуюся от первой характеристики загороживания, незагороженный элементарный пространственный сектор, имеющий первое расстояние до слушателя, и незагороженный элементарный пространственный сектор, имеющий второе расстояние до слушателя, при этом второе расстояние отличается от первого расстояния.

В дополнительных вариантах осуществления, модуль вычисления целевых данных выполнен с возможностью модификации или комбинирования частотно-зависимых

параметров дисперсии или ковариации в качестве элементов данных рендеринга для получения в качестве полного комбинированного результата полной комбинированной дисперсии или полного комбинированного параметра ковариации, и вычисления по меньшей мере одного из сигнальной метки интерауральной или межканальной когерентности, сигнальной метки интерауральной или межканальной разности уровней, сигнальной метки интерауральной или межканальной разности фаз, первого бокового усиления или второго бокового усиления в качестве целевых данных рендеринга.

В дополнительных вариантах осуществления предусмотрен модуль инициализации для определения по меньшей мере одного из левого элемента данных дисперсии, правого элемента данных дисперсии и элемента данных ковариации из предварительно сохраненных данных передаточной функции восприятия звука человеком, при этом модуль инициализации выполнен с возможностью вычисления левого элемента данных дисперсии, правого элемента данных дисперсии или элемента данных ковариации из множества данных передаточной функции восприятия звука человеком для ограниченного пространственного сектора, и при этом ограниченный пространственный сектор имеет такой размер, что по меньшей мере два элемента левых данных передаточной функции восприятия звука человеком по меньшей мере два элемента правых данных передаточной функции восприятия звука человеком существуют для ограниченного пространственного диапазона.

Список литературы

Alary, B., Politis, A. и Välimäki, V., 2017 год, "Velvet Noise Decorrelator".

Baumgarte, F. и Faller, C., 2003 год, "Binaural Cue Coding-Part I: Psychoacoustic Fundamentals and Design Principles", *Speech and Audio Processing, IEEE Transactions on*, 11(6), стр. 509-519.

Blauert, J., 2001 год, *Spatial hearing (3 Ausg.)*. Cambridge; Mass: MIT Press.

Faller, C. и Baumgarte, F., 2003 год, "Binaural Cue Coding-Part II: Schemes and Applications", *Speech and Audio Processing, IEEE Transactions on*, 11(6), стр. 520-531.

Kendall, G. S., 1995 год, "The Decorrelation of Audio Signals and Its Impact on Spatial Imagery", *Computer Music Journal*, 19(4), стр. 71-87.

Lauridsen, H., 1954 год, "Experiments Concerning Different Kinds of Room-Acoustics Recording", *Ingenioren*, 47.

Pihlajamäki, T., Santala, O. и Pulkki, V., 2014 год, "Synthesis of Spatially Extended Virtual Source with Time-Frequency Decomposition of Mono Signals", *Journal of the Audio Engineering Society*, 62(7/8), стр. 467-484.

Potard, G., 2003 год, "The study on sound source apparent shape and wideness".

Potard, G. и Burnett, I., 2004 год, "Decorrelation Techniques for the Rendering of Apparent Sound Source Width in 3D Audio Displays".

Pulkki, V., 1997 год, "Virtual Sound Source Positioning Using Vector Base Amplitude Panning", *Journal of the Audio Engineering Society*, 45(6), стр. 456-466.

Pulkki, V., 1999 год, "Uniform spreading of amplitude panned virtual sources".

Pulkki, V., 2007 год, "Spatial Sound Reproduction with Directional Audio Coding", *J. Audio Eng. Soc.*, 55(6), стр. 503-516.

Pulkki, V., Laitinen, M.-V. и Erkut, C., 2009 год, "Efficient Spatial Sound Synthesis for Virtual Worlds".

Schlecht, S. J., Alary, B., Välimäki, V. и Habets, E. A., 2018 год, "Optimized Velvet-Noise Decorrelator".

Schmele, T. и Sayin, U., 2018 год, "Controlling the Apparent Source Size in Ambisonics Using Decorrelation Filters".

Schmidt, J. и Schröder, E. F., 2004 год, "New and Advanced Features for Audio Presentation in the MPEG-4 Standard".

Verron, C., Aramaki, M., Kronland-Martinet, R. и Pallone, G., 2010 год, "The 3-D Immersive Synthesizer for Environmental Sounds", Audio, Speech and Language Processing, IEEE Transactions on, title="A Backward-Compatible Multichannel Audio Codec", 18(6), стр. 1550-1561.

Zotter, F. и Frank, M., 2013 год, "Efficient Phantom Source Widening", Archives of Acoustics, 38(1), стр. 27-37.

Zotter, F., Frank, M., Kronlachner, M. и Choi, J.-W., 2014 год, "Efficient Phantom Source Widening and Diffuseness in Ambisonics".

(57) Формула изобретения

1. Устройство для синтезирования пространственно протяженного источника (7000) звука (SESS), содержащее:

15 устройство (200, 2000) хранения данных для сохранения элементов данных рендеринга для различных элементарных пространственных секторов, покрывающих диапазон рендеринга для слушателя;

процессор (4000) идентификации секторов для идентификации, из различных элементарных пространственных секторов, набора элементарных пространственных секторов, принадлежащих пространственно протяженному источнику звука, на основе 20 данных слушателя и данных пространственно протяженных источников звука, при этом набор элементарных пространственных секторов содержит два или более элементарных пространственных сектора из различных элементарных пространственных секторов;

25 модуль (5000) вычисления целевых данных для вычисления целевых данных рендеринга с использованием сочетания элементов данных рендеринга для набора элементарных пространственных секторов; и

аудиопроцессор (300, 3000) для обработки аудиосигнала, представляющего пространственно протяженный источник звука, с использованием целевых данных 30 рендеринга.

2. Устройство по п. 1, в котором устройство (200, 2000) хранения данных выполнено с возможностью сохранения (810) в качестве элементов данных рендеринга для каждого элементарного пространственного сектора по меньшей мере одного из левого элемента данных дисперсии, связанного с левыми данными передаточной функции восприятия 35 звука, правого элемента данных дисперсии, связанного с правыми данными передаточной функции восприятия звука (HRTF), и элемента данных ковариации, связанного с левыми данными HRTF и правыми данными HRTF,

при этом модуль (5000) вычисления целевых данных выполнен с возможностью суммирования (830) левых элементов данных дисперсии для набора элементарных пространственных секторов или правых элементов данных дисперсии для набора 40 элементарных пространственных секторов, или элементов данных ковариации для набора элементарных пространственных секторов, соответственно, для получения по меньшей мере одного суммированного элемента,

при этом модуль (5000) вычисления целевых данных выполнен с возможностью вычисления (840) по меньшей мере одной сигнальной метки рендеринга в качестве 45 целевых данных рендеринга по меньшей мере из одного суммированного элемента, и

при этом аудиопроцессор (300, 3000) выполнен с возможностью обработки (850) аудиосигнала с использованием по меньшей мере одной сигнальной метки рендеринга.

3. Устройство по п. 1 или 2, в котором процессор (4000) идентификации секторов выполнен с возможностью применения алгоритма проецирования или анализа по трассировке лучей для определения набора элементарных пространственных секторов, или использования в качестве данных слушателя положения слушателя или ориентации слушателя либо использования в качестве данных пространственно протяженных источников звука (SESS) ориентации SESS, положение SESS или информации относительно геометрии SESS.

4. Устройство по одному из предшествующих пунктов, в котором процессор (4000) идентификации секторов выполнен с возможностью:

приёма из описания аудиосцены информации загораживания относительно потенциально загораживающего объекта (7010), и

определения на основе информации загораживания конкретного пространственного сектора из набора элементарных пространственных секторов в качестве загораживающего сектора, и

при этом модуль (5000) вычисления целевых данных выполнен с возможностью применения (5020) функции загораживания к элементам данных рендеринга, сохраненным для загораживающего сектора, таким образом, чтобы получить модифицированные данные и использовать модифицированные данные для вычисления (5060) целевых данных рендеринга.

5. Устройство по п. 4, в котором функция загораживания представляет собой функцию нижних частот, имеющую различные значения ослабления для различных частот, и при этом элементы данных рендеринга представляют собой элементы данных для различных частот, и при этом модуль (5000) вычисления целевых данных выполнен с возможностью взвешивания (5020) для нескольких частот элемента данных для определенной частоты со значением ослабления для определенной частоты для получения модифицированных данных рендеринга.

6. Устройство по п. 4 или 5, в котором процессор (4000) идентификации секторов выполнен с возможностью определения (4010) того, что другой элементарный пространственный сектор из набора элементарных пространственных секторов, определенных для загораживающего объекта, не загораживается посредством потенциального загораживающего объекта, и при этом модуль (5000) вычисления целевых данных выполнен с возможностью комбинирования (5040) модифицированных данных из загораживающего сектора и элементов данных рендеринга другого сектора без модификации с использованием функции загораживания либо с модификацией посредством другой функции модификации для получения целевых данных рендеринга.

7. Устройство по одному из предшествующих пунктов, в котором процессор (4000) идентификации секторов выполнен с возможностью определения первого элементарного пространственного сектора из набора элементарных пространственных секторов как имеющего первую характеристику и определения второго элементарного

пространственного сектора из набора элементарных пространственных секторов как имеющего вторую отличающуюся характеристику, и при этом модуль (5000) вычисления целевых данных выполнен с возможностью неприменения (4010) функции модификации к первому элементарному пространственному сектору и применения (4020) функции модификации ко второму элементарному пространственному сектору либо применения (4020) первой функцию модификации к первому элементарному пространственному сектору и применения (4030) второй функции модификации ко второму элементарному пространственному сектору, причем вторая функция модификации отличается от первой функции модификации.

8. Устройство по п. 7, в котором первая функция модификации является частотно-избирательной, и вторая функция модификации является постоянной по частоте, либо при этом первая функция модификации имеет первую частотно-избирательную характеристику, и при этом вторая функция модификации имеет вторую частотно-избирательную характеристику, отличающуюся от первой частотно-избирательной характеристики, либо при этом первая функция модификации имеет первую характеристику ослабления, и вторая функция модификации имеет вторую отличающуюся характеристику ослабления, и при этом модуль (5000) вычисления целевых данных выполнен с возможностью выбора или регулировки функции модификации из первой функции модификации и второй функции модификации на основе расстояния между первым элементарным пространственным сектором или вторым элементарным пространственным сектором и слушателем либо на основе характеристики объекта, расположенного между слушателем и соответствующим элементарным пространственным сектором.

9. Устройство по одному из предшествующих пунктов, в котором процессор (4000) идентификации секторов выполнен с возможностью классификации набора элементарных пространственных секторов на различные классы секторов на основе характеристик, ассоциированных с элементарными пространственными секторами, при этом модуль (5000) вычисления целевых данных выполнен с возможностью комбинирования (5020) элементов данных рендеринга элементарных пространственных секторов в каждом классе для получения комбинированного результата для каждого класса, если более одного элементарного пространственного сектора находится в классе, и применения конкретной функции модификации, ассоциированной по меньшей мере с одним классом, к комбинированному результату этого класса для получения модифицированного комбинированного результата для этого класса, или применения конкретной функции модификации, ассоциированной по меньшей мере с одним классом, к одному или более элементов данных одного или более элементарных пространственных секторов каждого класса для получения модифицированных элементов данных и комбинирования модифицированных элементов данных элементарных пространственных секторов в каждом классе для получения модифицированного комбинированного результата для этого класса, комбинирования (5040) комбинированного результата, либо, при наличии, модифицированного комбинированного результата для каждого класса для получения полного комбинированного результата и использования полного комбинированного результата в качестве целевых данных рендеринга или вычисления (5060) целевых данных рендеринга из полного комбинированного результата.

10. Устройство по п. 9, в котором характеристика для элементарного пространственного сектора определяется как представляющий собой одно из группы, содержащей загороженный элементарный пространственный сектор, включающий в себя первую характеристику загороживания, загороженный элементарный пространственный сектор, включающий в себя вторую характеристику загороживания, отличающуюся от первой характеристики загороживания, незагороженный элементарный пространственный сектор, имеющий первое расстояние до слушателя, и незагороженный элементарный пространственный сектор, имеющий второе расстояние до слушателя, при этом второе расстояние отличается от первого расстояния.

11. Устройство по п. 9 или 10, в котором модуль (5000) вычисления целевых данных выполнен с возможностью модификации или комбинирования (5020, 5040) частотно-

зависимых параметров дисперсии или ковариации в качестве элементов данных рендеринга для получения в качестве полного комбинированного результата полной комбинированной дисперсии или полного комбинированного параметра ковариации, и вычисления (5060) по меньшей мере одного из сигнальной метки интерауральной когерентности, сигнальной метки интерауральной разности уровней, сигнальной метки интерауральной разности фаз, первого бокового усиления или второго бокового усиления в качестве целевых данных рендеринга.

12. Устройство по одному из предшествующих пунктов, в котором аудиопроцессор (300, 3000) выполнен с возможностью выполнения по меньшей мере одного из регулирования (320, 3200) межканальной когерентности, регулирования (330, 3300) межканальной разности фаз, регулирования (340, 3400) межканальной разности уровней с использованием соответствующих сигнальных меток в качестве целевых данных рендеринга.

13. Устройство по одному из предшествующих пунктов, в котором диапазон рендеринга содержит сферу или часть сферы вокруг слушателя, при этом диапазон рендеринга привязывается к положению слушателя или к ориентации слушателя, и при этом каждый элементарный пространственный сектор имеет размер азимута и размер подъема.

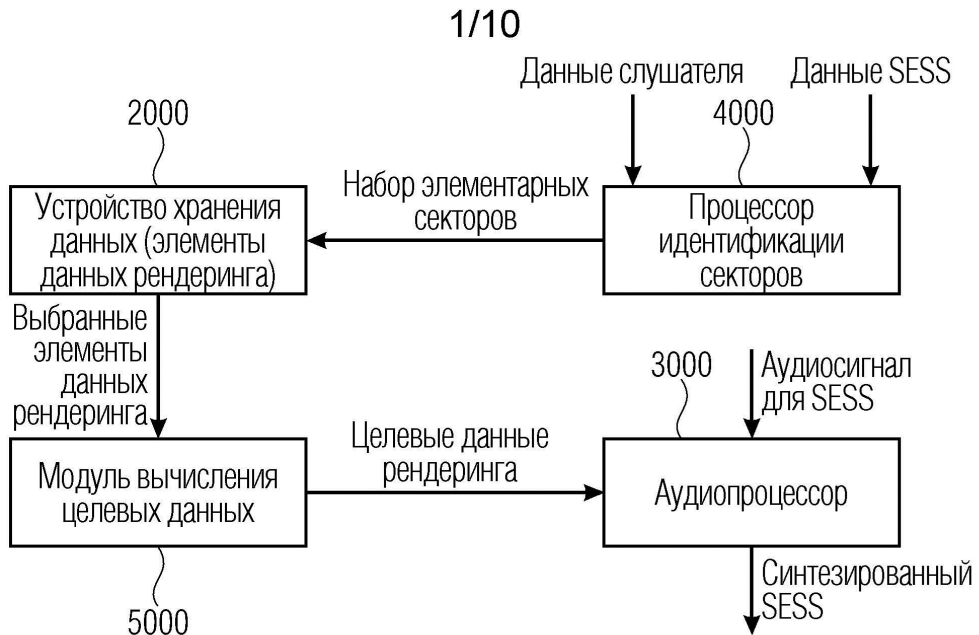
14. Устройство по п. 13, в котором размер азимута и размер подъема элементарных пространственных секторов отличаются друг от друга, так что размер азимута является более точным для элементарного пространственного сектора непосредственно перед слушателем по сравнению с размером азимута элементарного пространственного сектора больше вбок от слушателя, либо при этом размер азимута снижается в направлении вбок от слушателя, либо при этом размер подъема элементарного пространственного сектора меньше размера азимута этого сектора.

15. Способ синтеза пространственно протяженного источника звука (SESS), содержащий этапы, на которых:

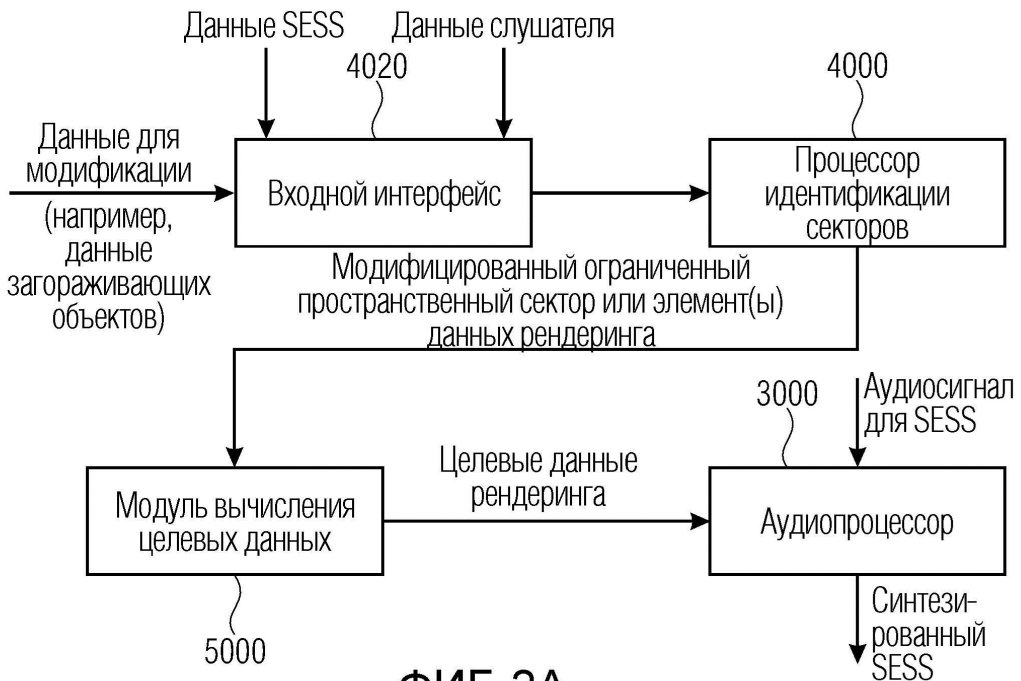
сохраняют элементы данных рендеринга для различных элементарных пространственных секторов, покрывающих диапазон рендеринга для слушателя; идентифицируют из различных элементарных пространственных секторов набор элементарных пространственных секторов, принадлежащих пространственно протяженному источнику звука, на основе данных слушателя и данных пространственно протяженных источников звука, при этом набор элементарных пространственных секторов содержит два или более элементарных пространственных сектора из различных элементарных пространственных секторов;

вычисляют целевые данные рендеринга с использованием сочетания элементов данных рендеринга для набора элементарных пространственных секторов; и

обрабатывают аудиосигнал, представляющий пространственно протяженный источник звука, с использованием целевых данных рендеринга.

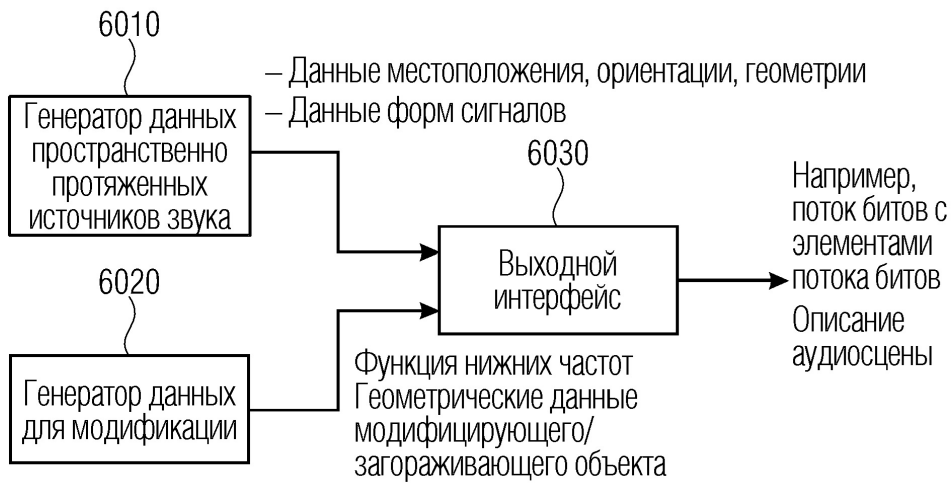


ФИГ. 1
(ПЕРВЫЙ АСПЕКТ)

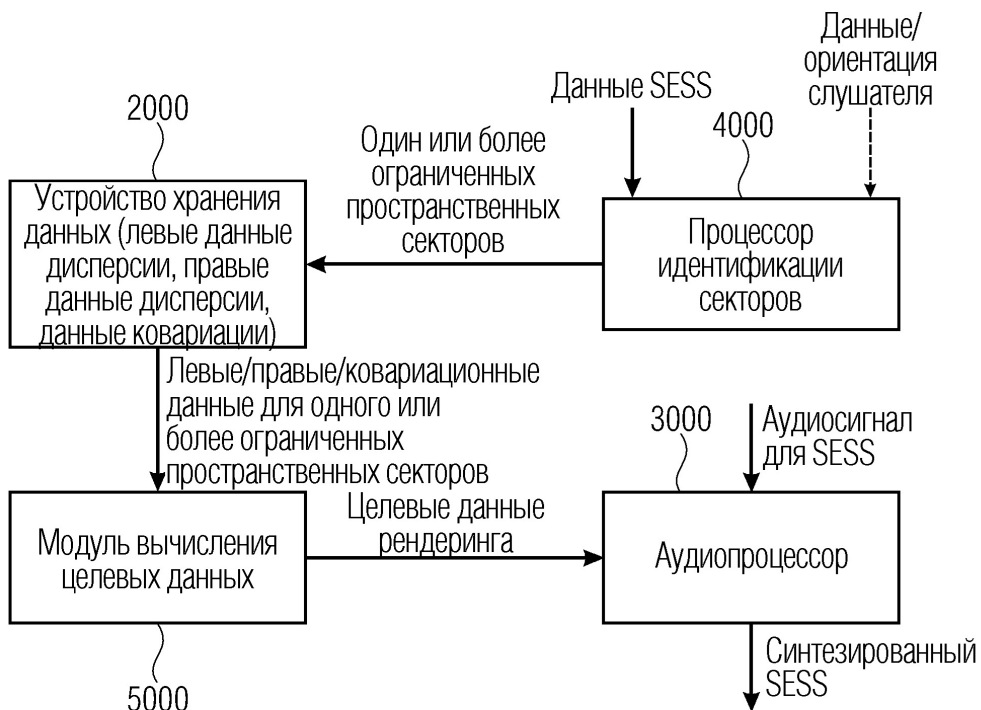


ФИГ. 2А
(ВТОРОЙ АСПЕКТ)

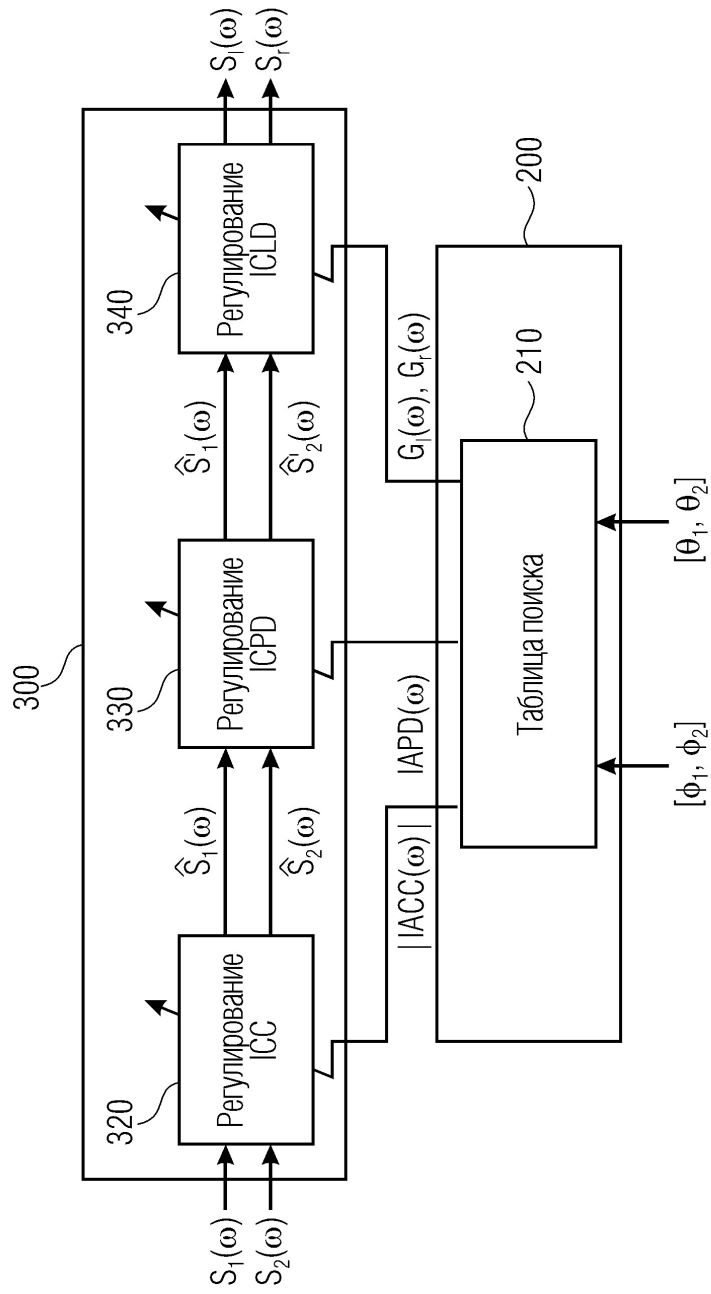
2/10



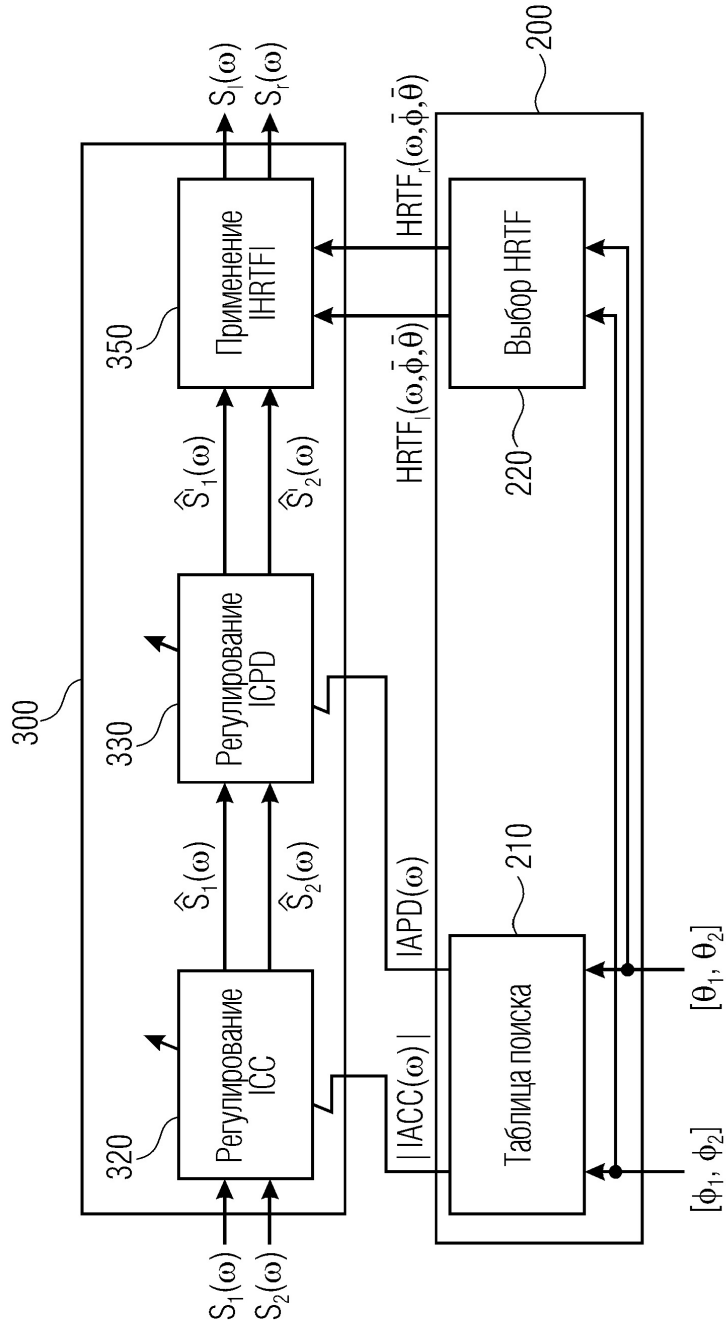
ФИГ. 2В
(ВТОРОЙ АСПЕКТ)



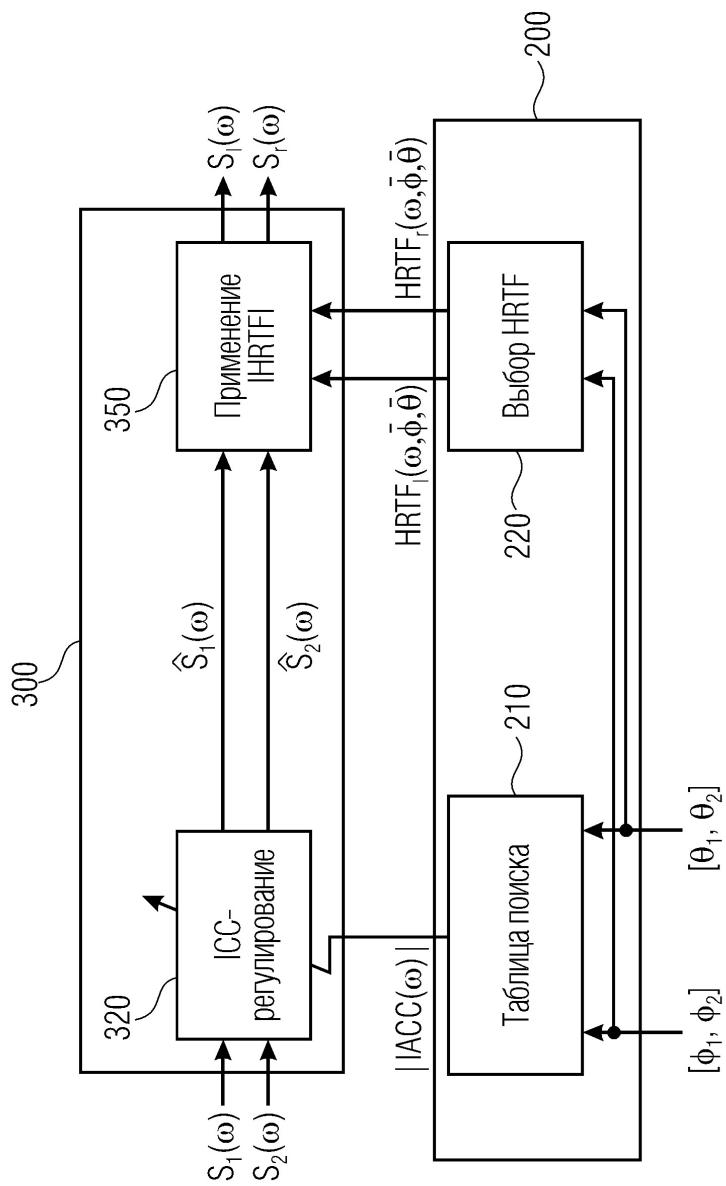
ФИГ. 3
(ТРЕТИЙ АСПЕКТ)



ФИГ. 4

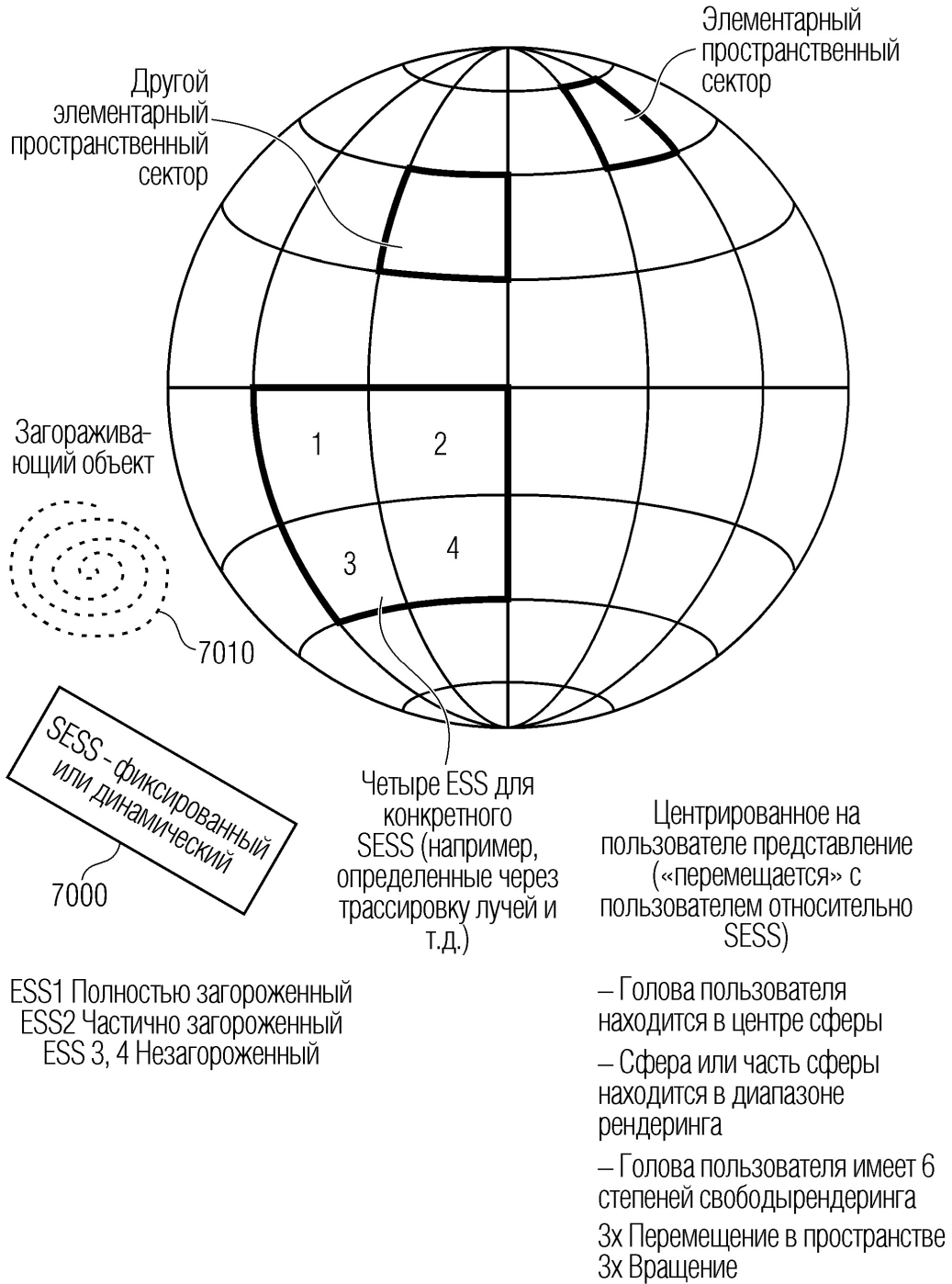


ФИГ. 5



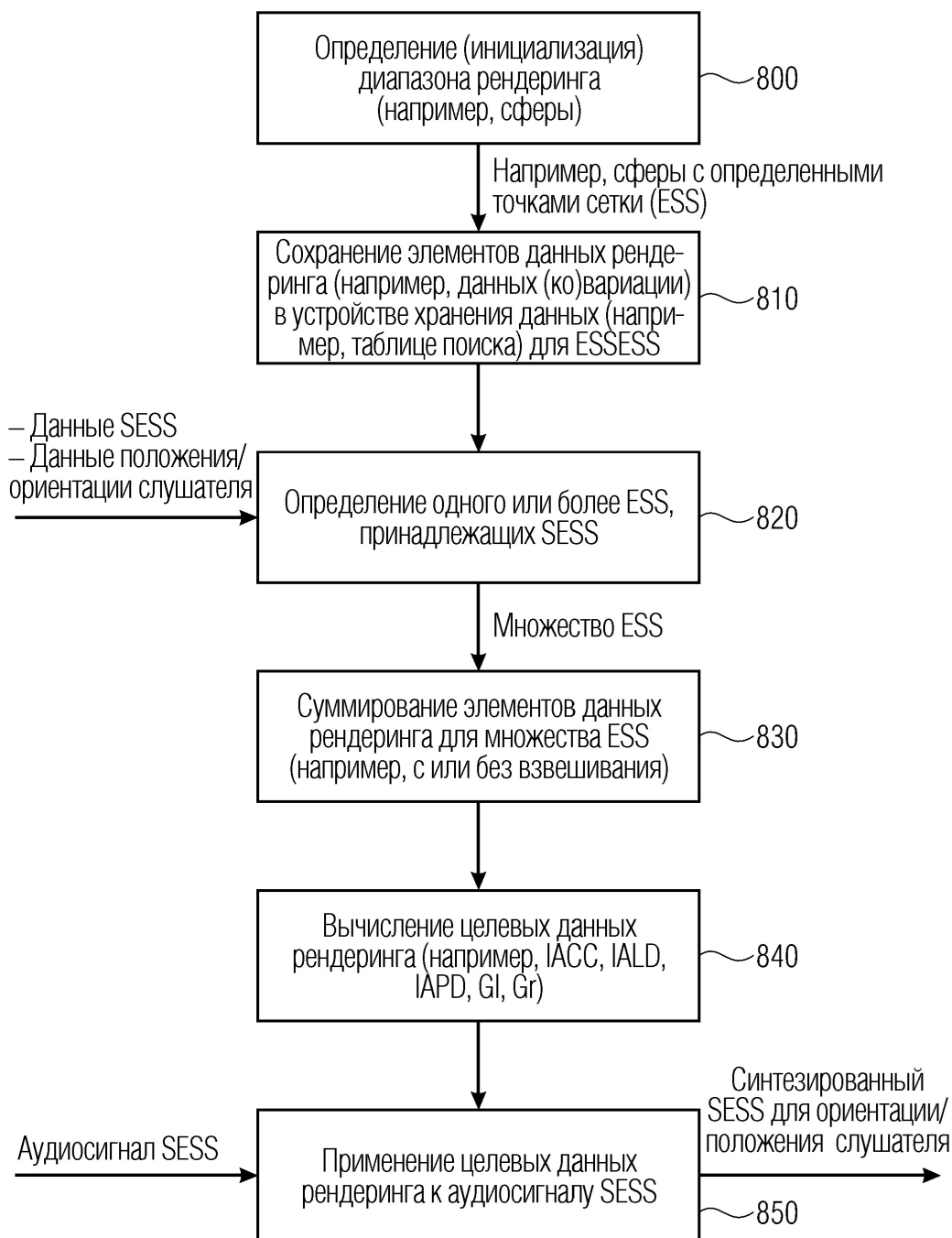
ФИГ. 6

6/10



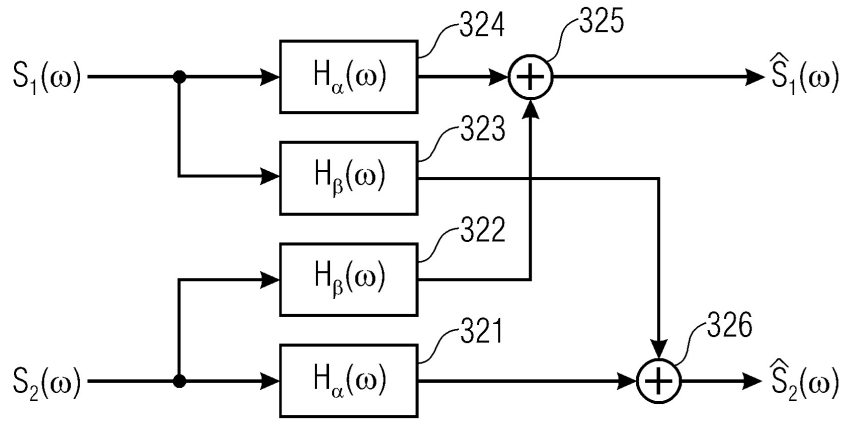
ФИГ. 7

7/10

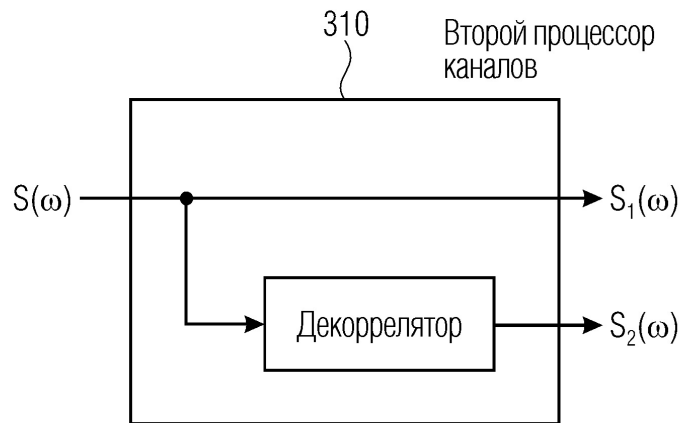


ФИГ. 8

8/10

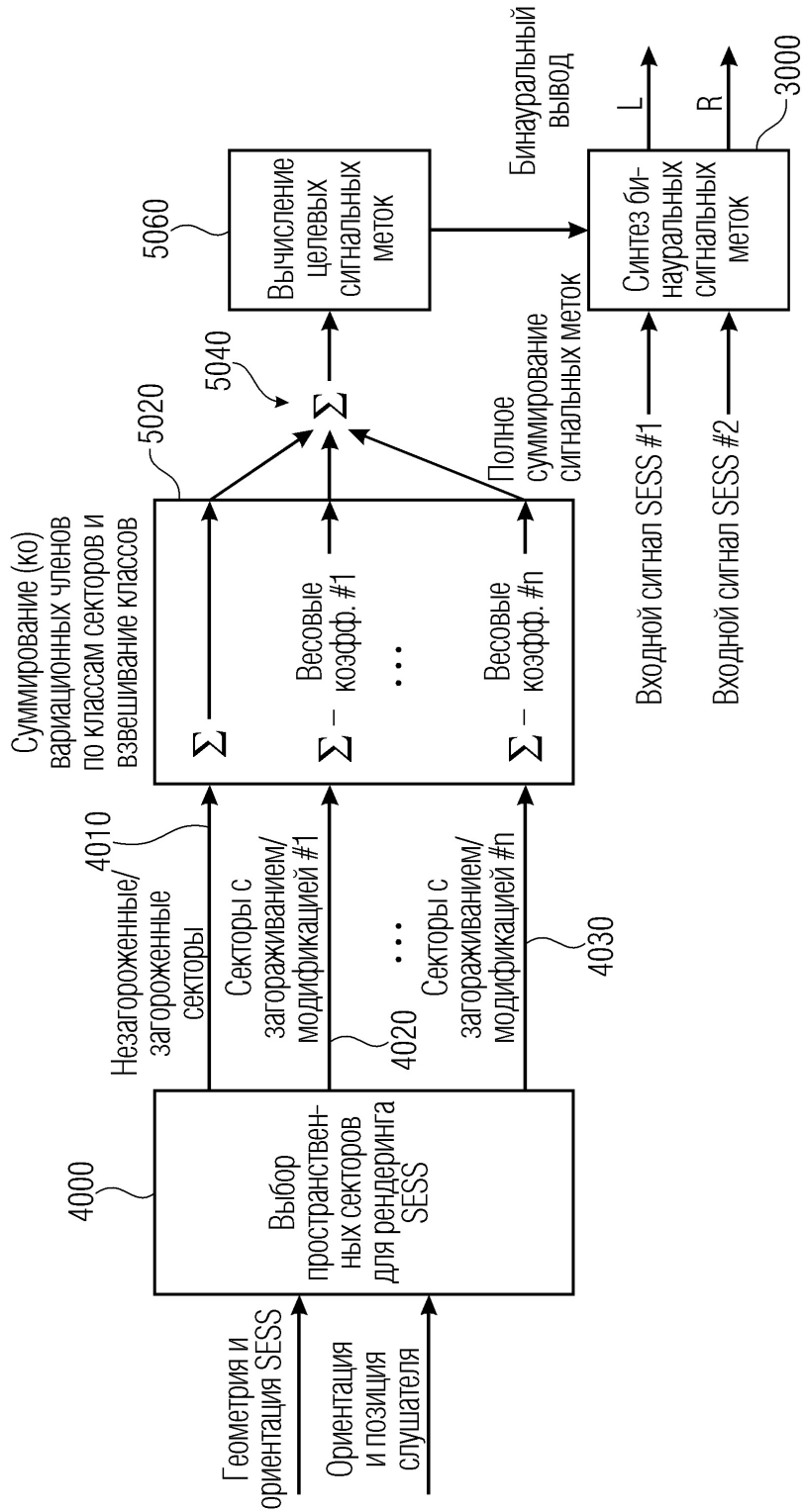


ФИГ. 9



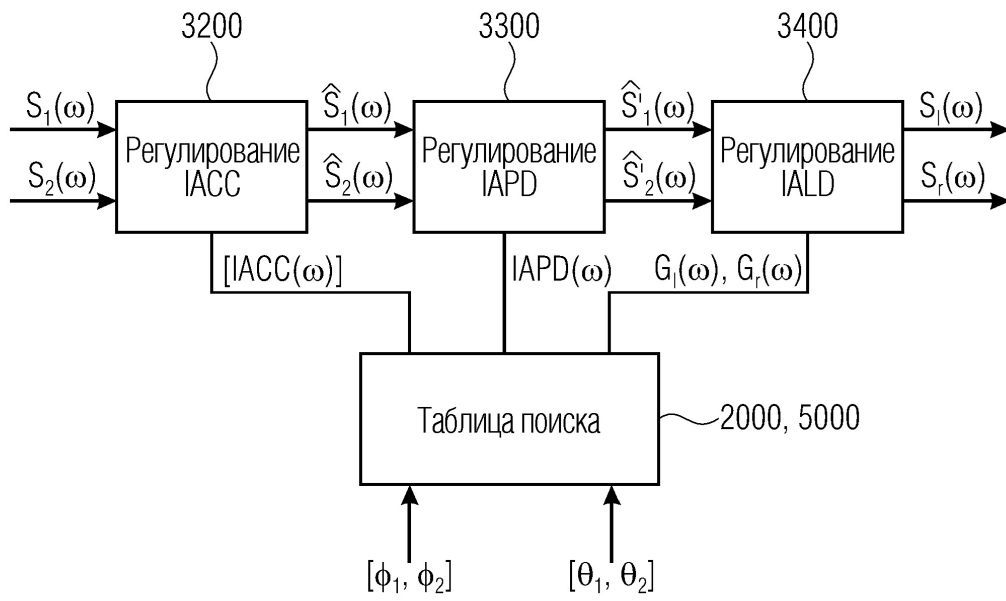
ФИГ. 10

9/10

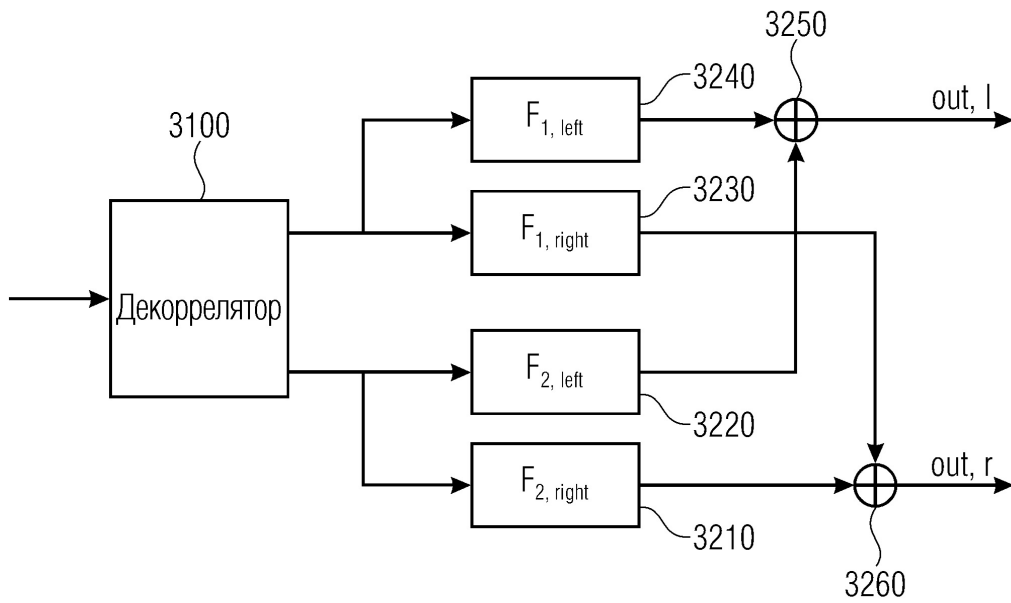


ФИГ. 11

10/10



ФИГ. 12



ФИГ. 13