



ФЕДЕРАЛЬНАЯ СЛУЖБА
ПО ИНТЕЛЛЕКТУАЛЬНОЙ СОБСТВЕННОСТИ

(12) ОПИСАНИЕ ИЗОБРЕТЕНИЯ К ПАТЕНТУ

(52) СПК

H04S 7/00 (2006.01); H04R 25/00 (2006.01); H04R 25/407 (2006.01)

(21)(22) Заявка: 2016147370, 23.04.2015

(24) Дата начала отсчета срока действия патента:
23.04.2015

Дата регистрации:
03.08.2018

Приоритет(ы):

(30) Конвенционный приоритет:
05.05.2014 EP 14167053.9;
05.09.2014 EP 14183854.0

(43) Дата публикации заявки: 06.06.2018 Бюл. № 16

(45) Опубликовано: 03.08.2018 Бюл. № 22

(85) Дата начала рассмотрения заявки РСТ на
национальной фазе: 05.12.2016

(86) Заявка РСТ:
EP 2015/058857 (23.04.2015)

(87) Публикация заявки РСТ:
WO 2015/169617 (12.11.2015)

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

(72) Автор(ы):

ХАБЕТС Эмануэль (DE),
ТИРГАРТ Оливер (DE),
КОВАЛЬЧИК Конрад (DE)

(73) Патентообладатель(и):

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

(56) Список документов, цитированных в отчете
о поиске: EP 2346028 A1, 20.07.2011. EP
2600343 A1, 05.06.2013. WO 2006/014449 A1,
09.02.2006. US 7583805 B2, 01.09.2009. WO
2012/033950 A1, 15.03.2012. RU 2416172 C1,
10.04.2011. RU 2363116 C2, 27.07.2009.

(54) СИСТЕМА, УСТРОЙСТВО И СПОСОБ ДЛЯ СОВМЕСТИМОГО ВОСПРОИЗВЕДЕНИЯ
АКУСТИЧЕСКОЙ СЦЕНЫ НА ОСНОВЕ АДАПТИВНЫХ ФУНКЦИЙ

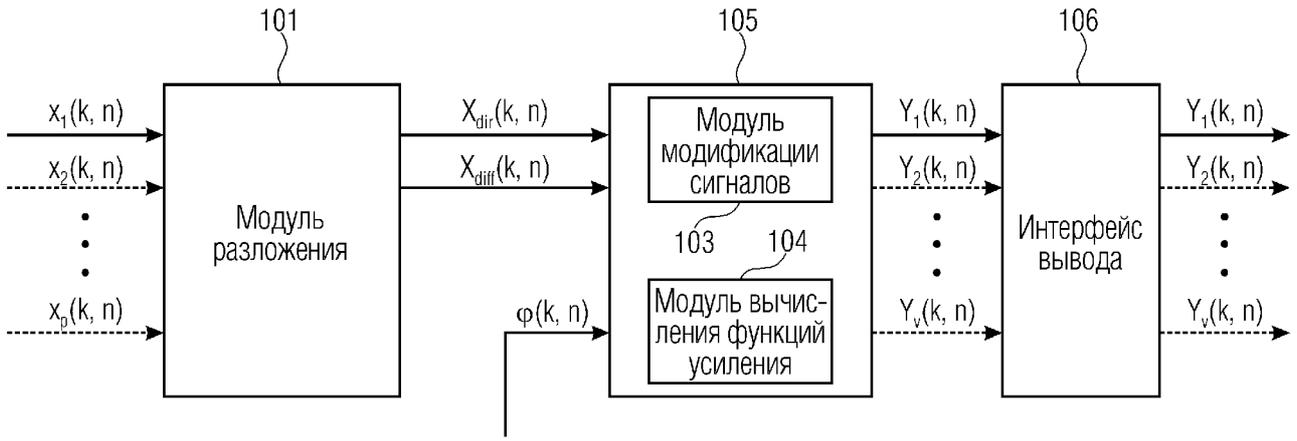
(57) Реферат:

Изобретение относится к средствам для воспроизведения акустической сцены. Технический результат заключается в повышении эффективности обработки аудиосигнала. Принимают сигнал прямых компонент, содержащий компоненты прямых сигналов из двух или более исходных аудиосигналов. Принимают сигнал диффузных компонент, содержащий компоненты диффузных сигналов из упомянутых двух или более исходных аудиосигналов. Принимают информацию

направления, при этом упомянутая информация направления зависит от направления прибытия компонент прямых сигналов из упомянутых двух или более исходных аудиосигналов. Генерируют один или более обработанных диффузных сигналов в зависимости от сигнала диффузных компонент. Для каждого выходного аудиосигнала из упомянутых одного или более выходных аудиосигналов определяют в зависимости от направления прибытия усиление прямого звука, применяют упомянутое усиление прямого звука

к сигналу прямых компонент, чтобы получить
обработанный прямой сигнал. 3 н. и 14 з.п. ф-лы,

21 ил.



ФИГ. 1С

RU 2663343 C2

RU 2663343 C2



FEDERAL SERVICE
FOR INTELLECTUAL PROPERTY

(51) Int. Cl.
H04S 7/00 (2006.01)
H04R 25/00 (2006.01)
G10L 19/008 (2013.01)

(12) ABSTRACT OF INVENTION

(52) CPC

H04S 7/00 (2006.01); H04R 25/00 (2006.01); H04R 25/407 (2006.01)(21)(22) Application: **2016147370, 23.04.2015**(24) Effective date for property rights:
23.04.2015Registration date:
03.08.2018

Priority:

(30) Convention priority:
05.05.2014 EP 14167053.9;
05.09.2014 EP 14183854.0(43) Application published: **06.06.2018** Bull. № 16(45) Date of publication: **03.08.2018** Bull. № 22(85) Commencement of national phase: **05.12.2016**(86) PCT application:
EP 2015/058857 (23.04.2015)(87) PCT publication:
WO 2015/169617 (12.11.2015)Mail address:
129090, Moskva, ul. B. Spasskaya, 25, str. 3, OOO
"Yuridicheskaya firma Gorodisskij i Partnery"

(72) Inventor(s):

HABETS, Emanuel (DE),
THIERGART, Oliver (DE),
KOWALCZYK, Konrad (DE)

(73) Proprietor(s):

FRAUNHOFER-GESELLSCHAFT ZUR
FOERDERUNG DER ANGEWANDTEN
FORSCHUNG e.V. (DE)**(54) SYSTEM, DEVICE AND METHOD FOR COMPATIBLE REPRODUCTION OF ACOUSTIC SCENE
BASED ON ADAPTIVE FUNCTIONS**

(57) Abstract:

FIELD: acoustics.

SUBSTANCE: invention relates to means for reproducing an acoustic scene. In the method, a direct component signal is received that contains components of direct signals from two or more original audio signals. Also, a signal of diffuse components comprising components of diffuse signals from said two or more original audio signals is received. And the direction information is received, wherein said direction information depends on the direction of arrival of the direct signal components from the two or more original

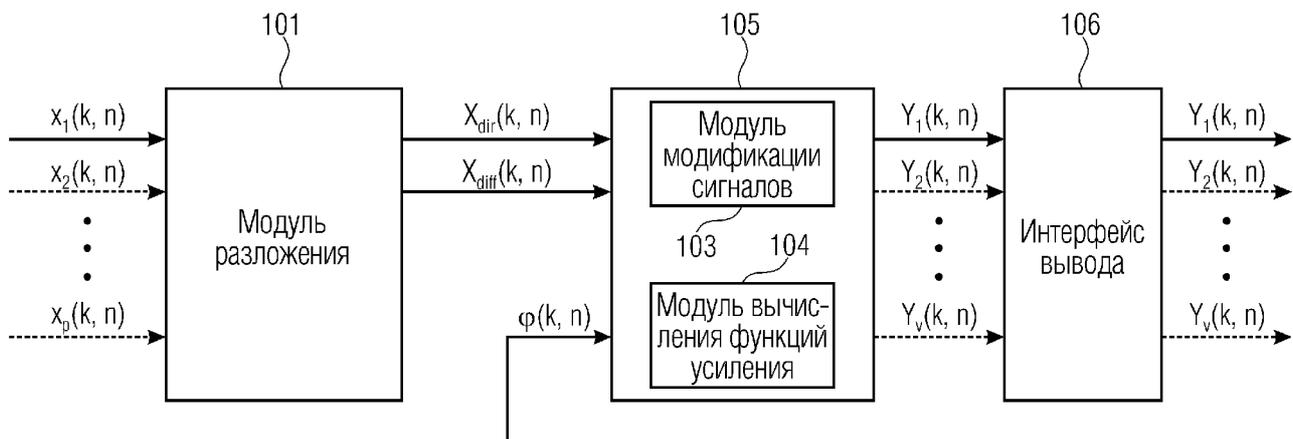
audio signals. One or more processed diffuse signals are generated depending on the signal of diffuse components. For each output audio signal from said one or more output audio signals, a direct sound gain is determined depending on the direction of arrival, said amplification of direct sound is applied to the signal of direct components to obtain a processed direct signal.

EFFECT: technical result is to improve the efficiency of audio signal processing.

17 cl, 21 dwg

RU 2 663 343 C2

RU 2 663 343 C2



ФИГ. 1С

RU 2663343 C2

RU 2663343 C2

Настоящее изобретение относится к обработке аудиосигналов, и, в частности, к системе, устройству и способу для совместимого воспроизведения акустической сцены на основе информированной пространственной фильтрации.

В воспроизведении пространственного звука звук в местоположении записи (стороне ближнего конца) захватывается с помощью множества микрофонов и затем воспроизводится на стороне воспроизведения (стороне дальнего конца) с использованием множества громкоговорителей или наушников. Во многих применениях, требуется воспроизводить записанный звук таким образом, чтобы пространственное изображение, воссоздаваемое на стороне дальнего конца, было совместимо с исходным пространственным изображением на стороне ближнего конца. Это означает, например, что звук источников звука воспроизводится из направлений, где источники присутствовали в исходном сценарии записи. Альтернативно, когда, например, видео дополняет записанное аудио, является желательным, чтобы звук воспроизводился таким образом, чтобы воссоздаваемое акустическое изображение было совместимо с видеоизображением. Это означает, например, что звук источника звука воспроизводится из направления, где источник является видимым на видео. Дополнительно, видеокамера может оснащаться функцией визуального масштабирования или пользователь на стороне дальнего конца может применять цифровое масштабирование к видео, которое изменяет визуальное изображение. В этом случае, акустическое изображение воспроизводимого пространственного звука должно изменяться соответствующим образом. Во многих случаях, сторона дальнего конца определяет пространственное изображение, с которым воспроизводимый звук должен быть совместимым, при этом оно определяется либо на стороне дальнего конца, либо во время проигрывания, например, когда вовлечено видеоизображение. Следовательно, пространственный звук на стороне ближнего конца должен записываться, обрабатываться, и передаваться таким образом, чтобы на стороне дальнего конца мы все еще могли управлять воссоздаваемым акустическим изображением.

Возможность воспроизводить записанную акустическую сцену совместимым образом с требуемым пространственным изображением требуется во многих современных применениях. Например, современные пользовательские устройства, такие как цифровые камеры или мобильные телефоны, часто оснащены видеокамерой и множеством микрофонов. Это обеспечивает возможность записывать видео вместе с пространственным звуком, например, стереозвуком. При воспроизведении записанного аудио вместе с видео, требуется, чтобы визуальное и акустическое изображение были совместимыми. Когда пользователь увеличивает масштаб с помощью камеры, является желательным воссоздавать эффект визуального масштабирования акустически, чтобы визуальное и акустическое изображения были выровнены при просмотре видео. Например, когда пользователь увеличивает масштаб на человеке, речь этого человека должна становиться менее реверберационной по мере того, как человек показывается более близко к камере. Более того, речь человека должна воспроизводиться из того же направления, где человек показывается в визуальном изображении. Имитация визуального масштабирования камеры акустически в последующем упоминается как акустическое масштабирование и представляет один пример совместимого воспроизведения аудио-видео. Совместимое воспроизведение аудио-видео, которое может включать в себя акустическое масштабирование, также является полезным в телеконференцсвязи, где пространственный звук на стороне ближнего конца воспроизводится на стороне дальнего конца вместе с визуальным изображением. Более того, является желательным воссоздавать эффект визуального масштабирования

акустически, чтобы визуальное и акустическое изображения были выровнены.

Первый вариант осуществления акустического масштабирования был представлен в [1], где эффект масштабирования получается посредством увеличения направленности направленного микрофона второго порядка, чей сигнал генерируется на основе сигналов

линейного массива микрофонов. Этот подход был расширен в [2] до стерео масштабирования. Более недавний подход для моно или стерео масштабирования был представлен в [3], который состоит в изменении уровней источников звука таким образом, чтобы источник из фронтального направления сохранялся, тогда как источники, приходящие из других направлений, и диффузный звук ослаблялись.

Подходы, предложенные в [1, 2], дают результатом увеличение отношения прямого звука к реверберации (DRR) и подход в [3] дополнительно обеспечивает возможность для подавления нежелательных источников. Вышеупомянутые подходы предполагают, что источник звука располагается спереди камеры, и не имеют целью захватывать акустическое изображение, которое является совместимым с видеоизображением.

Хорошо известный подход для гибких записи и воспроизведения пространственного звука представлен посредством направленного аудио кодирования (DirAC) [4]. В DirAC, пространственный звук на стороне ближнего конца описывается исходя из аудиосигнала и параметрической вспомогательной информации, именно направления прибытия (DOA) и диффузности звука. Параметрическое описание обеспечивает возможность воспроизведения исходного пространственного изображения с произвольными установками громкоговорителей. Это означает, что воссоздаваемое пространственное изображение на стороне дальнего конца является совместимым с пространственным изображением во время записи на стороне ближнего конца. Однако, если, например, видео дополняет записанное аудио, то воспроизводимый пространственный звук не необходимо выровнен с видеоизображением. Более того, воссоздаваемое акустическое изображение не может регулироваться, когда визуальное изображение изменяется, например, когда направление просмотра и масштабирование камеры изменяется. Это означает, что DirAC не обеспечивает никакой возможности регулировать воссоздаваемое акустическое изображение для произвольного требуемого пространственного изображения.

В [5], акустическое масштабирование реализовано на основе DirAC. DirAC представляет разумную основу для реализации акустического масштабирования, так как оно основывается на простой, но все же мощной модели сигналов, предполагающей, что звуковое поле в частотно-временной области состоит из одиночной плоской волны плюс диффузный звук. Лежащие в основе параметры модели, например, DOA и диффузность, используются, чтобы разделять прямой звук и диффузный звук и создавать эффект акустического масштабирования. Параметрическое описание пространственного звука обеспечивает возможность эффективной передачи звуковой сцены стороне дальнего конца, при этом еще обеспечивает пользователя полным управлением над эффектом масштабирования и воспроизведением пространственного звука. Даже хотя DirAC использует множество микрофонов, чтобы оценивать параметры модели, применяются только одноканальные фильтры, чтобы извлекать прямой звук и диффузный звук, что ограничивает качество воспроизводимого звука. Более того, предполагается, что все источники в звуковой сцене располагаются на окружности и воспроизведение пространственного звука выполняется со ссылкой на изменяющееся положение аудиовизуальной камеры, что является несовместимым с визуальным масштабированием. Фактически, масштабирование изменяет угол обзора камеры, в то время как расстояние до визуальных объектов и их относительные положения в

изображении остаются неизменными, что отличается от перемещения камеры.

Родственный подход является так называемым способом виртуальных микрофонов (VM) [6,7], который учитывает такую же модель сигналов как DigAC, но обеспечивает возможность синтезировать сигнал несуществующего (виртуального) микрофона в произвольном положении в звуковой сцене. Перемещение VM в направлении к источнику звука является аналогичным перемещению камеры в новое положение. VM реализуется с использованием многоканальных фильтров, чтобы улучшать качество звука, но требует несколько распределенных массивов микрофонов, чтобы оценивать параметры модели.

Однако было бы весьма предпочтительным, если бы были обеспечены дополнительно улучшенные концепции для обработки аудиосигналов.

Таким образом, цель настоящего изобретения состоит в том, чтобы обеспечить улучшенные концепции для обработки аудиосигналов. Цель настоящего изобретения решается посредством системы согласно пункту 1 формулы изобретения, посредством устройства согласно пункту 14 формулы изобретения, посредством способа согласно пункту 15 формулы изобретения, посредством способа согласно пункту 16 формулы изобретения и посредством компьютерной программы согласно пункту 17 формулы изобретения.

Обеспечивается система для генерирования одного или более выходных аудиосигналов. Система содержит модуль генерирования сигнала компонент, сигнальный процессор, и интерфейс вывода. Модуль генерирования сигнала компонент сконфигурирован с возможностью принимать два или более входных аудиосигналов, при этом модуль генерирования сигнала компонент сконфигурирован с возможностью генерировать сигнал прямых компонент, содержащий компоненты прямых сигналов упомянутых двух или более входных аудиосигналов, и при этом модуль генерирования сигнала компонент сконфигурирован с возможностью генерировать сигнал диффузных компонент, содержащий компоненты диффузных сигналов упомянутых двух или более входных аудиосигналов. Сигнальный процессор сконфигурирован с возможностью принимать сигнал прямых компонент, сигнал диффузных компонент и информацию направления, при этом упомянутая информация направления зависит от направления прибытия компонент прямых сигналов упомянутых двух или более входных аудиосигналов. Более того, сигнальный процессор сконфигурирован с возможностью генерировать один или более обработанных диффузных сигналов в зависимости от сигнала диффузных компонент. Для каждого выходного аудиосигнала из упомянутых одного или более выходных аудиосигналов, сигнальный процессор сконфигурирован с возможностью определять, в зависимости от направления прибытия, усиление прямого звука, сигнальный процессор сконфигурирован с возможностью применять упомянутое усиление прямого звука к сигналу прямых компонент, чтобы получать обработанный прямой сигнал, и сигнальный процессор сконфигурирован с возможностью комбинировать упомянутый обработанный прямой сигнал и один из упомянутых одного или более обработанных диффузных сигналов, чтобы генерировать упомянутый выходной аудиосигнал. Интерфейс вывода сконфигурирован с возможностью выводить упомянутые один или более выходных аудиосигналов. Сигнальный процессор содержит модуль вычисления функций усиления для вычисления одной или более функций усиления, при этом каждая функция усиления из упомянутых одной или более функций усиления содержит множество значений аргумента функции усиления, при этом возвращаемое значение функции усиления назначено каждому из упомянутых значений аргумента функции усиления, при этом, когда упомянутая функция усиления принимает

одно из упомянутых значений аргумента функции усиления, упомянутая функция усиления сконфигурирована с возможностью возвращать возвращаемое значение функции усиления, которое назначено упомянутому одному из упомянутых значений аргумента функции усиления. Более того, сигнальный процессор дополнительно
5 содержит модуль модификации сигналов для выбора, в зависимости от направления прибытия, зависящего от направления значения аргумента из значений аргумента функции усиления для функции усиления из упомянутых одной или более функций усиления, для получения возвращаемого значения функции усиления, которое назначено упомянутому зависящему от направления значению аргумента, от упомянутой функции
10 усиления, и для определения значения усиления, по меньшей мере, одного из упомянутых одного или более выходных аудиосигналов в зависимости от упомянутого возвращаемого значения функции усиления, полученного от упомянутой функции усиления.

Согласно одному варианту осуществления, модуль вычисления функций усиления
15 может, например, быть сконфигурирован с возможностью генерировать таблицу поиска для каждой функции усиления из упомянутых одной или более функций усиления, при этом таблица поиска содержит множество записей, при этом каждая из записей таблицы поиска содержит одно из значений аргумента функции усиления и возвращаемое
20 значение функции усиления, которое назначено упомянутому значению аргумента функции усиления, при этом модуль вычисления функций усиления может, например, быть сконфигурирован с возможностью хранить таблицу поиска каждой функции усиления в постоянной или непостоянной памяти, и при этом модуль модификации сигналов может, например, быть сконфигурирован с возможностью получать
25 возвращаемое значение функции усиления, которое назначено упомянутому зависящему от направления значению аргумента, посредством считывания упомянутого возвращаемого значения функции усиления из одной из упомянутых одной или более таблиц поиска, которые сохранены в памяти.

В одном варианте осуществления, сигнальный процессор может, например, быть сконфигурирован с возможностью определять два или более выходных аудиосигналов,
30 при этом модуль вычисления функций усиления может, например, быть сконфигурирован с возможностью вычислять две или более функции усиления, при этом, для каждого выходного аудиосигнала из упомянутых двух или более выходных аудиосигналов, модуль вычисления функций усиления может, например, быть сконфигурирован с
35 возможностью вычислять функцию усиления панорамирования, которая назначена упомянутому выходному аудиосигналу в качестве одной из упомянутых двух или более функций усиления, при этом модуль модификации сигналов может, например, быть сконфигурирован с возможностью генерировать упомянутый выходной аудиосигнал в зависимости от упомянутой функции усиления панорамирования.

Согласно одному варианту осуществления, функция усиления панорамирования
40 каждого из упомянутых двух или более выходных аудиосигналов может, например, иметь один или более глобальных максимумов, являющихся одним из значений аргумента функции усиления упомянутой функции усиления панорамирования, при этом для каждого из упомянутых одного или более глобальных максимумов упомянутой функции усиления панорамирования, не существует никакое другое значение аргумента
45 функции усиления, для которого упомянутая функция усиления панорамирования возвращает более большое возвращаемое значение функции усиления, чем для упомянутых глобальных максимумов, и при этом, для каждой пары из первого выходного аудиосигнала и второго выходного аудиосигнала из упомянутых двух или

более выходных аудиосигналов, по меньшей мере, один из упомянутых одного или более глобальных максимумов функции усиления панорамирования первого выходного аудиосигнала может, например, отличаться от любого из упомянутых одного или более глобальных максимумов функции усиления панорамирования второго выходного аудиосигнала.

5 Согласно одному варианту осуществления, для каждого выходного аудиосигнала из упомянутых двух или более выходных аудиосигналов, модуль вычисления функций усиления может, например, быть сконфигурирован с возможностью вычислять оконную функцию усиления, которая назначена упомянутому выходному аудиосигналу в качестве одной из упомянутых двух или более функций усиления, при этом модуль модификации сигналов может, например, быть сконфигурирован с возможностью генерировать упомянутый выходной аудиосигнал в зависимости от упомянутой оконной функции усиления, и при этом, если значение аргумента упомянутой оконной функции усиления больше, чем нижний порог окна и меньше, чем верхний порог окна, оконная функция усиления сконфигурирована с возможностью возвращать возвращаемое значение функции усиления, которое больше, чем любое возвращаемое значение функции усиления, возвращаемое упомянутой оконной функцией усиления, если значение аргумента оконной функции меньше, чем нижний порог, или больше, чем верхний порог.

10 В одном варианте осуществления, оконная функция усиления каждого из упомянутых двух или более выходных аудиосигналов имеет один или более глобальных максимумов, являющихся одним из значений аргумента функции усиления упомянутой оконной функции усиления, при этом для каждого из упомянутых одного или более глобальных максимумов упомянутой оконной функции усиления, не существует никакое другое значение аргумента функции усиления, для которого упомянутая оконная функция усиления возвращает более большое возвращаемое значение функции усиления, чем для упомянутых глобальных максимумов, и при этом, для каждой пары из первого выходного аудиосигнала и второго выходного аудиосигнала из упомянутых двух или более выходных аудиосигналов, по меньшей мере, один из упомянутых одного или более глобальных максимумов оконной функции усиления первого выходного аудиосигнала может, например, быть равным одному из упомянутых одного или более глобальных максимумов оконной функции усиления второго выходного аудиосигнала.

15 Согласно одному варианту осуществления, модуль вычисления функций усиления может, например, быть сконфигурирован с возможностью дополнительно принимать информацию ориентации, указывающую угловой сдвиг направления просмотра по отношению к направлению прибытия, и при этом модуль вычисления функций усиления может, например, быть сконфигурирован с возможностью генерировать функцию усиления панорамирования каждого из выходных аудиосигналов в зависимости от информации ориентации.

20 В одном варианте осуществления, модуль вычисления функций усиления может, например, быть сконфигурирован с возможностью генерировать оконную функцию усиления каждого из выходных аудиосигналов в зависимости от информации ориентации.

25 Согласно одному варианту осуществления, модуль вычисления функций усиления может, например, быть сконфигурирован с возможностью дополнительно принимать информацию масштабирования, при этом информация масштабирования указывает угол раскрытия камеры, и при этом модуль вычисления функций усиления может, например, быть сконфигурирован с возможностью генерировать функцию усиления панорамирования каждого из выходных аудиосигналов в зависимости от информации

масштабирования.

В одном варианте осуществления, модуль вычисления функций усиления может, например, быть сконфигурирован с возможностью генерировать оконную функцию усиления каждого из выходных аудиосигналов в зависимости от информации

5 масштабирования.

Согласно одному варианту осуществления, модуль вычисления функций усиления может, например, быть сконфигурирован с возможностью дополнительно принимать параметр калибровки для выравнивания визуального изображения и акустического изображения, и при этом модуль вычисления функций усиления может, например, быть

10 сконфигурирован с возможностью генерировать функцию усиления панорамирования каждого из выходных аудиосигналов в зависимости от параметра калибровки.

В одном варианте осуществления, модуль вычисления функций усиления может, например, быть сконфигурирован с возможностью генерировать оконную функцию усиления каждого из выходных аудиосигналов в зависимости от параметра калибровки.

15 Система согласно одному из предшествующих утверждений, в которой модуль вычисления функций усиления может, например, быть сконфигурирован с возможностью принимать информацию о визуальном изображении, и модуль вычисления функций усиления может, например, быть сконфигурирован с возможностью генерировать, в зависимости от информации о визуальном изображении, функцию размытия,

20 возвращающую комплексные усиления, чтобы реализовать перцепционное рассеивание источника звука.

Более того, обеспечивается устройство для генерирования одного или более выходных аудиосигналов. Устройство содержит сигнальный процессор и интерфейс вывода.

Сигнальный процессор сконфигурирован с возможностью принимать сигнал прямых компонент, содержащий компоненты прямых сигналов из упомянутых двух или более

25 исходных аудиосигналов, при этом сигнальный процессор сконфигурирован с возможностью принимать сигнал диффузных компонент, содержащий компоненты диффузных сигналов из упомянутых двух или более исходных аудиосигналов, и при этом сигнальный процессор сконфигурирован с возможностью принимать информацию

30 направления, при этом упомянутая информация направления зависит от направления прибытия компонент прямых сигналов упомянутых двух или более входных аудиосигналов. Более того, сигнальный процессор сконфигурирован с возможностью генерировать один или более обработанных диффузных сигналов в зависимости от

сигнала диффузных компонент. Для каждого выходного аудиосигнала из упомянутых

35 одного или более выходных аудиосигналов, сигнальный процессор сконфигурирован с возможностью определять, в зависимости от направления прибытия, усиление прямого звука, сигнальный процессор сконфигурирован с возможностью применять упомянутое усиление прямого звука к сигналу прямых компонент, чтобы получать обработанный прямой сигнал, и сигнальный процессор сконфигурирован с возможностью

40 комбинировать упомянутый обработанный прямой сигнал и один из упомянутых одного или более обработанных диффузных сигналов, чтобы генерировать упомянутый выходной аудиосигнал. Интерфейс вывода сконфигурирован с возможностью выводить упомянутые один или более выходных аудиосигналов. Сигнальный процессор содержит модуль вычисления функций усиления для вычисления одной или более функций

45 усиления, при этом каждая функция усиления из упомянутых одной или более функций усиления содержит множество значений аргумента функции усиления, при этом возвращаемое значение функции усиления назначено каждому из упомянутых значений аргумента функции усиления, при этом, когда упомянутая функция усиления принимает

одно из упомянутых значений аргумента функции усиления, упомянутая функция усиления сконфигурирована с возможностью возвращать возвращаемое значение функции усиления, которое назначено упомянутому одному из упомянутых значений аргумента функции усиления. Более того, сигнальный процессор дополнительно
5 содержит модуль модификации сигналов для выбора, в зависимости от направления прибытия, зависящего от направления значения аргумента из значений аргумента функции усиления для функции усиления из упомянутых одной или более функций усиления, для получения возвращаемого значения функции усиления, которое назначено упомянутому зависящему от направления значению аргумента, от упомянутой функции
10 усиления, и для определения значения усиления, по меньшей мере, одного из упомянутых одного или более выходных аудиосигналов в зависимости от упомянутого возвращаемого значения функции усиления, полученного от упомянутой функции усиления.

Дополнительно, обеспечивается способ для генерирования одного или более
15 выходных аудиосигналов. Способ содержит:

- Прием двух или более входных аудиосигналов.
- Генерирование сигнала прямых компонент, содержащего компоненты прямых сигналов упомянутых двух или более входных аудиосигналов.
- Генерирование сигнала диффузных компонент, содержащего компоненты
20 диффузных сигналов упомянутых двух или более входных аудиосигналов.
- Прием информации направления в зависимости от направления прибытия компонент прямых сигналов упомянутых двух или более входных аудиосигналов.
- Генерирование одного или более обработанных диффузных сигналов в зависимости от сигнала диффузных компонент.
- 25 - Для каждого выходного аудиосигнала из упомянутых одного или более выходных аудиосигналов, определение, в зависимости от направления прибытия, усиления прямого звука, применение упомянутого усиления прямого звука к сигналу прямых компонент, чтобы получать обработанный прямой сигнал, и комбинирование упомянутого обработанного прямого сигнала и одного из упомянутых одного или более
30 обработанных диффузных сигналов, чтобы генерировать упомянутый выходной аудиосигнал. И:

- Вывод упомянутых одного или более выходных аудиосигналов.

Генерирование упомянутых одного или более выходных аудиосигналов содержит вычисление одной или более функций усиления, при этом каждая функция усиления из
35 упомянутых одной или более функций усиления содержит множество значений аргумента функции усиления, при этом возвращаемое значение функции усиления назначено каждому из упомянутых значений аргумента функции усиления, при этом, когда упомянутая функция усиления принимает одно из упомянутых значений аргумента функции усиления, упомянутая функция усиления сконфигурирована с возможностью
40 возвращать возвращаемое значение функции усиления, которое назначено упомянутому одному из упомянутых значений аргумента функции усиления. Более того, генерирование упомянутых одного или более выходных аудиосигналов содержит выбор, в зависимости от направления прибытия, зависящего от направления значения аргумента из значений аргумента функции усиления для функции усиления из упомянутых одной или более
45 функций усиления, для получения возвращаемого значения функции усиления, которое назначено упомянутому зависящему от направления значению аргумента, от упомянутой функции усиления, и для определения значения усиления, по меньшей мере, одного из упомянутых одного или более выходных аудиосигналов в зависимости от упомянутого

возвращаемого значения функции усиления, полученного от упомянутой функции усиления.

Более того, обеспечивается способ для генерирования одного или более выходных аудиосигналов. Способ содержит:

- 5 - Прием сигнала прямых компонент, содержащего компоненты прямых сигналов из упомянутых двух или более исходных аудиосигналов.
 - прием сигнала диффузных компонент, содержащего компоненты диффузных сигналов из упомянутых двух или более исходных аудиосигналов.
 - Прием информации направления, при этом упомянутая информация направления
- 10 зависит от направления прибытия компонент прямых сигналов упомянутых двух или более входных аудиосигналов.
 - Генерирование одного или более обработанных диффузных сигналов в зависимости от сигнала диффузных компонент.
 - Для каждого выходного аудиосигнала из упомянутых одного или более выходных
- 15 аудиосигналов, определение, в зависимости от направления прибытия, усиления прямого звука, применение упомянутого усиления прямого звука к сигналу прямых компонент, чтобы получать обработанный прямой сигнал, и комбинирование упомянутого обработанного прямого сигнала и одного из упомянутых одного или более
- 20 обработанных диффузных сигналов, чтобы генерировать упомянутый выходной аудиосигнал. И:
 - Вывод упомянутых одного или более выходных аудиосигналов.

Генерирование упомянутых одного или более выходных аудиосигналов содержит вычисление одной или более функций усиления, при этом каждая функция усиления из упомянутых одной или более функций усиления содержит множество значений аргумента

25 функции усиления, при этом возвращаемое значение функции усиления назначено каждому из упомянутых значений аргумента функции усиления, при этом, когда упомянутая функция усиления принимает одно из упомянутых значений аргумента функции усиления, упомянутая функция усиления сконфигурирована с возможностью

30 возвращать возвращаемое значение функции усиления, которое назначено упомянутому одному из упомянутых значений аргумента функции усиления. Более того, генерирование упомянутых одного или более выходных аудиосигналов содержит выбор, в зависимости от направления прибытия, зависящего от направления значения аргумента из значений

35 аргумента функции усиления для функции усиления из упомянутых одной или более функций усиления, для получения возвращаемого значения функции усиления, которое назначено упомянутому зависящему от направления значению аргумента, от упомянутой функции усиления, и для определения значения усиления, по меньшей мере, одного из упомянутых одного или более выходных аудиосигналов в зависимости от упомянутого возвращаемого значения функции усиления, полученного от упомянутой функции усиления.

40 Более того, обеспечиваются компьютерные программы, при этом каждая из компьютерных программ сконфигурирована с возможностью осуществлять один из вышеописанных способов, когда исполняется на компьютере или сигнальном процессоре, так что каждый из вышеописанных способов осуществляется посредством одной из компьютерных программ.

45 Дополнительно, обеспечивается система для генерирования одного или более выходных аудиосигналов. Система содержит модуль генерирования сигнала компонент, сигнальный процессор, и интерфейс вывода. Модуль генерирования сигнала компонент сконфигурирован с возможностью принимать два или более входных аудиосигналов,

при этом модуль генерирования сигнала компонент сконфигурирован с возможностью генерировать сигнал прямых компонент, содержащий компоненты прямых сигналов упомянутых двух или более входных аудиосигналов, и при этом модуль генерирования сигнала компонент сконфигурирован с возможностью генерировать сигнал диффузных компонент, содержащий компоненты диффузных сигналов упомянутых двух или более входных аудиосигналов. Сигнальный процессор сконфигурирован с возможностью принимать сигнал прямых компонент, сигнал диффузных компонент и информацию направления, при этом упомянутая информация направления зависит от направления прибытия компонент прямых сигналов упомянутых двух или более входных аудиосигналов. Более того, сигнальный процессор сконфигурирован с возможностью генерировать один или более обработанных диффузных сигналов в зависимости от сигнала диффузных компонент. Для каждого выходного аудиосигнала из упомянутых одного или более выходных аудиосигналов, сигнальный процессор сконфигурирован с возможностью определять, в зависимости от направления прибытия, усиление прямого звука, сигнальный процессор сконфигурирован с возможностью применять упомянутое усиление прямого звука к сигналу прямых компонент, чтобы получать обработанный прямой сигнал, и сигнальный процессор сконфигурирован с возможностью комбинировать упомянутый обработанный прямой сигнал и один из упомянутых одного или более обработанных диффузных сигналов, чтобы генерировать упомянутый выходной аудиосигнал. Интерфейс вывода сконфигурирован с возможностью выводить упомянутые один или более выходных аудиосигналов.

Согласно вариантам осуществления, обеспечиваются концепции, чтобы достигать записи и воспроизведения пространственного звука таким образом, чтобы воссоздаваемое акустическое изображение могло, например, быть совместимым с требуемым пространственным изображением, которое, например, определяется пользователем на стороне дальнего конца или посредством видеоизображения. Предложенный подход использует массив микрофонов на стороне ближнего конца, который обеспечивает нам возможность разлагать захваченный звук на компоненты прямого звука и компоненту диффузного звука. Извлеченные компоненты звука затем передаются стороне дальнего конца. Совместимое воспроизведение пространственного звука может, например, реализовываться посредством взвешенной суммы извлеченных прямого звука и диффузного звука, где веса зависят от требуемого пространственного изображения, с которым воспроизводимый звук должен быть совместим, например, веса зависят от направления просмотра и коэффициента масштабирования видеокамеры, которые могут, например, дополнять запись аудио. Обеспечиваются концепции, которые используют информированные многоканальные фильтры для извлечения прямого звука и диффузного звука.

Согласно одному варианту осуществления, сигнальный процессор может, например, быть сконфигурирован с возможностью определять два или более выходных аудиосигналов, при этом для каждого выходного аудиосигнала из упомянутых двух или более выходных аудиосигналов функция усиления панорамирования может, например, быть назначена упомянутому выходному аудиосигналу, при этом функция усиления панорамирования каждого из упомянутых двух или более выходных аудиосигналов содержит множество значений аргумента функции панорамирования, при этом возвращаемое значение функции панорамирования может, например, быть назначено каждому из упомянутых значений аргумента функции панорамирования, при этом, когда упомянутая функция усиления панорамирования принимает одно из упомянутых значений аргумента функции панорамирования, упомянутая функция

усиления панорамирования может, например, быть сконфигурирована с возможностью возвращать возвращаемое значение функции панорамирования, которое назначено упомянутому одному из упомянутых значений аргумента функции панорамирования, и при этом сигнальный процессор может, например, быть сконфигурирован с
5 возможностью определять каждый из упомянутых двух или более выходных аудиосигналов в зависимости от зависящего от направления значения аргумента из значений аргумента функции панорамирования для функции усиления панорамирования, которая назначена упомянутому выходному аудиосигналу, при этом упомянутое зависящее от направления значение аргумента зависит от направления прибытия.

10 В одном варианте осуществления, функция усиления панорамирования каждого из упомянутых двух или более выходных аудиосигналов имеет один или более глобальных максимумов, являющихся одним из значений аргумента функции панорамирования, при этом для каждого из упомянутых одного или более глобальных максимумов каждой функции усиления панорамирования, не существует никакого другого значения аргумента
15 функции панорамирования, для которого упомянутая функция усиления панорамирования возвращает более большое возвращаемое значение функции панорамирования, чем для упомянутых глобальных максимумов, и при этом, для каждой пары из первого выходного аудиосигнала и второго выходного аудиосигнала из упомянутых двух или более выходных аудиосигналов, по меньшей мере, один из
20 упомянутых одного или более глобальных максимумов функции усиления панорамирования первого выходного аудиосигнала может, например, отличаться от любого из упомянутых одного или более глобальных максимумов функции усиления панорамирования второго выходного аудиосигнала.

Согласно одному варианту осуществления, сигнальный процессор может, например,
25 быть сконфигурирован с возможностью генерировать каждый выходной аудиосигнал из упомянутых одного или более выходных аудиосигналов в зависимости от оконной функции усиления, при этом оконная функция усиления может, например, быть сконфигурирована с возможностью возвращать возвращаемое значение оконной функции при приеме значения аргумента оконной функции, при этом, если значение
30 аргумента оконной функции может, например, быть больше, чем нижний порог окна и меньше, чем верхний порог окна, оконная функция усиления может, например, быть сконфигурирована с возможностью возвращать возвращаемое значение оконной функции, которое больше, чем любое возвращаемое значение оконной функции, возвращаемое оконной функцией усиления, если значение аргумента оконной функции
35 может, например, быть меньше, чем нижний порог, или больше, чем верхний порог.

В одном варианте осуществления, сигнальный процессор может, например, быть сконфигурирован с возможностью дополнительно принимать информацию ориентации, указывающую угловой сдвиг направления просмотра по отношению к направлению прибытия, и при этом, по меньшей мере, одна из функции усиления панорамирования
40 и оконной функции усиления зависит от информации ориентации; или при этом модуль вычисления функций усиления может, например, быть сконфигурирован с возможностью дополнительно принимать информацию масштабирования, при этом информация масштабирования указывает угол раскрытия камеры, и при этом, по меньшей мере, одна из функции усиления панорамирования и оконной функции усиления зависит от
45 информации масштабирования; или при этом модуль вычисления функций усиления может, например, быть сконфигурирован с возможностью дополнительно принимать параметр калибровки, и при этом, по меньшей мере, одна из функции усиления панорамирования и оконной функции усиления зависит от параметра калибровки.

Согласно одному варианту осуществления, сигнальный процессор может, например, быть сконфигурирован с возможностью принимать информацию расстояния, при этом сигнальный процессор может, например, быть сконфигурирован с возможностью генерировать каждый выходной аудиосигнал из упомянутых одного или более выходных аудиосигналов в зависимости от информации расстояния.

Согласно одному варианту осуществления, сигнальный процессор может, например, быть сконфигурирован с возможностью принимать исходное угловое значение в зависимости от исходного направления прибытия, которое является направлением прибытия компонент прямых сигналов упомянутых двух или более входных аудиосигналов, и может, например, быть сконфигурирован с возможностью принимать информацию расстояния, при этом сигнальный процессор может, например, быть сконфигурирован с возможностью вычислять модифицированное угловое значение в зависимости от исходного углового значения и в зависимости от информации расстояния, и при этом сигнальный процессор может, например, быть сконфигурирован с возможностью генерировать каждый выходной аудиосигнал из упомянутых одного или более выходных аудиосигналов в зависимости от модифицированного углового значения.

Согласно одному варианту осуществления, сигнальный процессор может, например, быть сконфигурирован с возможностью генерировать упомянутые один или более выходных аудиосигналов посредством выполнения низкочастотной фильтрации, или посредством добавления задержанного прямого звука, или посредством выполнения ослабления прямого звука, или посредством выполнения временного сглаживания, или посредством выполнения рассеивания направления прибытия, или посредством выполнения декорреляции.

В одном варианте осуществления, сигнальный процессор может, например, быть сконфигурирован с возможностью генерировать два или более выходных аудиоканалов, при этом сигнальный процессор может, например, быть сконфигурирован с возможностью применять усиление диффузного звука к сигналу диффузных компонент, чтобы получать промежуточный диффузный сигнал, и при этом сигнальный процессор может, например, быть сконфигурирован с возможностью генерировать один или более декоррелированных сигналов из промежуточного диффузного сигнала посредством выполнения декорреляции, при этом упомянутые один или более декоррелированных сигналов формируют упомянутые один или более обработанных диффузных сигналов, или при этом промежуточный диффузный сигнал и упомянутые один или более декоррелированных сигналов формируют упомянутые один или более обработанных диффузных сигналов.

Согласно одному варианту осуществления, сигнал прямых компонент и один или более дополнительных сигналов прямых компонент формируют группу из двух или более сигналов прямых компонент, при этом модуль генерирования сигнала компонент может, например, быть сконфигурирован с возможностью генерировать упомянутые один или более дополнительных сигналов прямых компонент, содержащих дополнительные компоненты прямых сигналов упомянутых двух или более входных аудиосигналов, при этом направление прибытия и одно или более дополнительных направления прибытий формируют группу из двух или более направлений прибытий, при этом каждое направление прибытия из группы из упомянутых двух или более направлений прибытий может, например, быть назначено в точности одному сигналу прямых компонент из группы из упомянутых двух или более сигналов прямых компонент, при этом количество сигналов прямых компонент из упомянутых двух или

более сигналов прямых компонент и количество направлений прибытий из упомянутых двух направлений прибытий могут, например, быть равными, при этом сигнальный процессор может, например, быть сконфигурирован с возможностью принимать группу из упомянутых двух или более сигналов прямых компонент, и группу из упомянутых 5 двух или более направлений прибытий, и при этом, для каждого выходного аудиосигнала из упомянутых одного или более выходных аудиосигналов, сигнальный процессор может, например, быть сконфигурирован с возможностью определять, для каждого сигнала прямых компонент из группы из упомянутых двух или более сигналов прямых компонент, усиление прямого звука в зависимости от направления прибытия 10 упомянутого сигнала прямых компонент, сигнальный процессор может, например, быть сконфигурирован с возможностью генерировать группу из двух или более обработанных прямых сигналов посредством применения, для каждого сигнала прямых компонент из группы из упомянутых двух или более сигналов прямых компонент, усиления прямого звука упомянутого сигнала прямых компонент к упомянутому 15 сигналу прямых компонент, и сигнальный процессор может, например, быть сконфигурирован с возможностью комбинировать один из упомянутых одного или более обработанных диффузных сигналов и каждый обработанный сигнал из группы из упомянутых двух или более обработанных сигналов, чтобы генерировать упомянутый выходной аудиосигнал.

20 В одном варианте осуществления, количество сигналов прямых компонент из группы из упомянутых двух или более сигналов прямых компонент плюс 1 может, например, быть меньше, чем количество входных аудиосигналов, которые принимаются интерфейсом приема.

Более того, может, например, обеспечиваться слуховой аппарат или вспомогательное 25 слуховое устройство, содержащее систему, как описана выше.

Более того, обеспечивается устройство для генерирования одного или более выходных аудиосигналов. Устройство содержит сигнальный процессор и интерфейс вывода. Сигнальный процессор сконфигурирован с возможностью принимать сигнал прямых компонент, содержащий компоненты прямых сигналов из упомянутых двух или более 30 исходных аудиосигналов, при этом сигнальный процессор сконфигурирован с возможностью принимать сигнал диффузных компонент, содержащий компоненты диффузных сигналов из упомянутых двух или более исходных аудиосигналов, и при этом сигнальный процессор сконфигурирован с возможностью принимать информацию направления, при этом упомянутая информация направления зависит от направления 35 прибытия компонент прямых сигналов упомянутых двух или более входных аудиосигналов. Более того, сигнальный процессор сконфигурирован с возможностью генерировать один или более обработанных диффузных сигналов в зависимости от сигнала диффузных компонент. Для каждого выходного аудиосигнала из упомянутых одного или более выходных аудиосигналов, сигнальный процессор сконфигурирован 40 с возможностью определять, в зависимости от направления прибытия, усиление прямого звука, сигнальный процессор сконфигурирован с возможностью применять упомянутое усиление прямого звука к сигналу прямых компонент, чтобы получать обработанный прямой сигнал, и сигнальный процессор сконфигурирован с возможностью комбинировать упомянутый обработанный прямой сигнал и один из упомянутых 45 одного или более обработанных диффузных сигналов, чтобы генерировать упомянутый выходной аудиосигнал. Интерфейс вывода сконфигурирован с возможностью выводить упомянутые один или более выходных аудиосигналов.

Дополнительно, обеспечивается способ для генерирования одного или более

выходных аудиосигналов. Способ содержит:

- Прием двух или более входных аудиосигналов.
- Генерирование сигнала прямых компонент, содержащего компоненты прямых сигналов упомянутых двух или более входных аудиосигналов.

5 - Генерирование сигнала диффузных компонент, содержащего компоненты диффузных сигналов упомянутых двух или более входных аудиосигналов.

- Прием информации направления в зависимости от направления прибытия компонент прямых сигналов упомянутых двух или более входных аудиосигналов.

10 - Генерирование одного или более обработанных диффузных сигналов в зависимости от сигнала диффузных компонент.

- для каждого выходного аудиосигнала из упомянутых одного или более выходных аудиосигналов, определение, в зависимости от направления прибытия, усиления прямого звука, применение упомянутого усиления прямого звука к сигналу прямых компонент, чтобы получать обработанный прямой сигнал, и комбинирование упомянутого
- 15 обработанного прямого сигнала и одного из упомянутых одного или более обработанных диффузных сигналов, чтобы генерировать упомянутый выходной аудиосигнал. И:

- Вывод упомянутых одного или более выходных аудиосигналов.

20 Более того, обеспечивается способ для генерирования одного или более выходных аудиосигналов. Способ содержит:

- Прием сигнала прямых компонент, содержащего компоненты прямых сигналов из упомянутых двух или более исходных аудиосигналов.

- Прием сигнала диффузных компонент, содержащего компоненты диффузных сигналов из упомянутых двух или более исходных аудиосигналов.

25 - Прием информации направления, при этом упомянутая информация направления зависит от направления прибытия компонент прямых сигналов упомянутых двух или более входных аудиосигналов.

- Генерирование одного или более обработанных диффузных сигналов в зависимости от сигнала диффузных компонент.

30 - Для каждого выходного аудиосигнала из упомянутых одного или более выходных аудиосигналов, определение, в зависимости от направления прибытия, усиления прямого звука, применение упомянутого усиления прямого звука к сигналу прямых компонент, чтобы получать обработанный прямой сигнал, и комбинирование упомянутого

35 обработанного прямого сигнала и одного из упомянутых одного или более обработанных диффузных сигналов, чтобы генерировать упомянутый выходной аудиосигнал. И:

- Вывод упомянутых одного или более выходных аудиосигналов.

40 Более того, обеспечиваются компьютерные программы, при этом каждая из компьютерных программ сконфигурирована с возможностью осуществлять один из вышеописанных способов, когда исполняется на компьютере или сигнальном процессоре, так что каждый из вышеописанных способов осуществляется посредством одной из компьютерных программ.

В последующем, варианты осуществления настоящего изобретения описываются более подробно со ссылкой на фигуры, на которых:

45 Фиг. 1a иллюстрирует систему согласно одному варианту осуществления, Фиг. 1b иллюстрирует устройство согласно одному варианту осуществления, Фиг. 1c иллюстрирует систему согласно другому варианту осуществления, Фиг. 1d иллюстрирует устройство согласно другому варианту осуществления,

Фиг. 2 показывает систему согласно другому варианту осуществления,

Фиг. 3 изображает модули генерирования сигнала компонент для разложения на прямой/диффузный звук и для параметра оценки системы согласно одному варианту осуществления,

5 Фиг. 4 показывает первую геометрию для воспроизведения акустической сцены с акустическим масштабированием согласно одному варианту осуществления, при этом источник звука располагается на фокальной плоскости,

Фиг. 5 иллюстрирует функции панорамирования для совместимого воспроизведения сцены и для акустического масштабирования,

10 Фиг. 6 изображает дополнительные функции панорамирования для совместимого воспроизведения сцены и для акустического масштабирования согласно вариантам осуществления,

Фиг. 7 иллюстрирует примерные оконные функции усиления для различных ситуаций согласно вариантам осуществления,

15 Фиг. 8 показывает функцию усиления диффузного звука согласно одному варианту осуществления,

Фиг. 9 изображает вторую геометрию для воспроизведения акустической сцены с акустическим масштабированием согласно одному варианту осуществления, при этом источник звука не располагается на фокальной плоскости,

20 Фиг. 10 иллюстрирует функции для описания размытия прямого звука, и

Фиг. 11 визуализирует слуховые аппараты согласно вариантам осуществления.

Фиг. 1а иллюстрирует систему для генерирования одного или более выходных аудиосигналов. Система содержит модуль 101 генерирования сигнала компонент, сигнальный процессор 105, и интерфейс 106 вывода.

25 Модуль 101 генерирования сигнала компонент сконфигурирован с возможностью генерировать сигнал прямых компонент $X_{\text{dir}}(k, n)$, содержащий компоненты прямых сигналов упомянутых двух или более входных аудиосигналов $x_1(k, n), x_2(k, n), \dots, x_p(k, n)$. Более того, модуль 101 генерирования сигнала компонент сконфигурирован с
30 возможностью генерировать сигнал диффузных компонент $X_{\text{diff}}(k, n)$, содержащий компоненты диффузных сигналов упомянутых двух или более входных аудиосигналов $x_1(k, n), x_2(k, n), \dots, x_p(k, n)$.

Сигнальный процессор 105 сконфигурирован с возможностью принимать сигнал прямых компонент $X_{\text{dir}}(k, n)$, сигнал диффузных компонент $X_{\text{diff}}(k, n)$ и информацию
35 направления, при этом упомянутая информация направления зависит от направления прибытия компонент прямых сигналов упомянутых двух или более входных аудиосигналов $x_1(k, n), x_2(k, n), \dots, x_p(k, n)$.

Более того, сигнальный процессор 105 сконфигурирован с возможностью генерировать один или более обработанных диффузных сигналов $Y_{\text{diff},1}(k, n), Y_{\text{diff},2}(k, n), \dots, Y_{\text{diff},v}(k, n)$ в зависимости от сигнала диффузных компонент $X_{\text{diff}}(k, n)$.
40

Для каждого выходного аудиосигнала $Y_i(k, n)$ из упомянутых одного или более выходных аудиосигналов $Y_1(k, n), Y_2(k, n), \dots, Y_v(k, n)$, сигнальный процессор 105 сконфигурирован с возможностью определять, в зависимости от направления прибытия,
45 усиление прямого звука $G_i(k, n)$, сигнальный процессор 105 сконфигурирован с возможностью применять упомянутое усиление прямого звука $G_i(k, n)$ к сигналу прямых компонент $X_{\text{dir}}(k, n)$, чтобы получать обработанный прямой сигнал $Y_{\text{dir},i}(k, n)$, и

сигнальный процессор 105 сконфигурирован с возможностью комбинировать упомянутый обработанный прямой сигнал $Y_{dir,i}(k, n)$ и один $Y_{diff,i}(k, n)$ из упомянутых одного или более обработанных диффузных сигналов $Y_{diff,1}(k, n), Y_{diff,2}(k, n), \dots, Y_{diff,v}(k, n)$, чтобы генерировать упомянутый выходной аудиосигнал $Y_i(k, n)$.

Интерфейс 106 вывода сконфигурирован с возможностью выводить упомянутые один или более выходных аудиосигналов $Y_1(k, n), Y_2(k, n), \dots, Y_v(k, n)$.

Как описано, информация направления зависит от направления прибытия $\phi(k, n)$ компонент прямых сигналов упомянутых двух или более входных аудиосигналов $x_1(k, n), x_2(k, n), \dots, x_p(k, n)$. Например, направление прибытия компонент прямых сигналов упомянутых двух или более входных аудиосигналов $x_1(k, n), x_2(k, n), \dots, x_p(k, n)$ может, например, само быть информацией направления. Или, например, информация направления, может, например, быть направлением распространения компонент прямых сигналов упомянутых двух или более входных аудиосигналов $x_1(k, n), x_2(k, n), \dots, x_p(k, n)$. В то время как направление прибытия обращено от массива микрофонов приема к источнику звука, направление распространения обращено от источника звука к массиву микрофонов приема. Таким образом, направление распространения обращено в точности в противоположном направлении по отношению к направлению прибытия и, поэтому, зависит от направления прибытия.

Чтобы генерировать один $Y_i(k, n)$ из упомянутых одного или более выходных аудиосигналов $Y_1(k, n), Y_2(k, n), \dots, Y_v(k, n)$, сигнальный процессор 105

- определяет, в зависимости от направления прибытия, усиление прямого звука $G_i(k, n)$,
- применяет упомянутое усиление прямого звука $G_i(k, n)$ к сигналу прямых компонент $X_{dir}(k, n)$, чтобы получать обработанный прямой сигнал $Y_{dir,i}(k, n)$, и

- комбинирует упомянутый обработанный прямой сигнал $Y_{dir,i}(k, n)$ и один $Y_{diff,i}(k, n)$ из упомянутых одного или более обработанных диффузных сигналов $Y_{diff,1}(k, n), Y_{diff,2}(k, n), \dots, Y_{diff,v}(k, n)$, чтобы генерировать упомянутый выходной аудиосигнал $Y_i(k, n)$

Это осуществляется для каждого из упомянутых одного или более выходных аудиосигналов $Y_1(k, n), Y_2(k, n), \dots, Y_v(k, n)$, которые должны генерироваться $Y_1(k, n), Y_2(k, n), \dots, Y_v(k, n)$. Сигнальный процессор может, например, быть сконфигурирован с возможностью генерировать один, два, три или более выходных аудиосигналов $Y_1(k, n), Y_2(k, n), \dots, Y_v(k, n)$.

Относительно упомянутых одного или более обработанных диффузных сигналов $Y_{diff,1}(k, n), Y_{diff,2}(k, n), \dots, Y_{diff,v}(k, n)$, согласно одному варианту осуществления, сигнальный процессор 105 может, например, быть сконфигурирован с возможностью генерировать упомянутые один или более обработанных диффузных сигналов $Y_{diff,1}(k, n), Y_{diff,2}(k, n), \dots, Y_{diff,v}(k, n)$ посредством применения усиления диффузного звука $Q(k, n)$ к сигналу диффузных компонент $X_{diff}(k, n)$.

Модуль 101 генерирования сигнала компонент может, например, быть сконфигурирован с возможностью генерировать сигнал прямых компонент $X_{dir}(k, n)$, содержащий компоненты прямых сигналов упомянутых двух или более входных

аудиосигналов $x_1(k, n)$, $x_2(k, n)$, ..., $x_p(k, n)$, и сигнал диффузных компонент $X_{\text{diff}}(k, n)$, содержащий компоненты диффузных сигналов упомянутых двух или более входных аудиосигналов $x_1(k, n)$, $x_2(k, n)$, ..., $x_p(k, n)$, посредством разложения упомянутых одного или более входных аудиосигналов на сигнал прямых компонент и на сигнал диффузных компонент.

В одном конкретном варианте осуществления, сигнальный процессор 105 может, например, быть сконфигурирован с возможностью генерировать два или более выходных аудиоканалов $Y_1(k, n)$, $Y_2(k, n)$, ..., $Y_v(k, n)$. Сигнальный процессор 105 может, например, быть сконфигурирован с возможностью применять усиление диффузного звука $Q(k, n)$ к сигналу диффузных компонент $X_{\text{diff}}(k, n)$, чтобы получать промежуточный диффузный сигнал. Более того, сигнальный процессор 105 может, например, быть сконфигурирован с возможностью генерировать один или более декоррелированных сигналов из промежуточного диффузного сигнала посредством выполнения декорреляции, при этом упомянутые один или более декоррелированных сигналов формируют упомянутые один или более обработанных диффузных сигналов $Y_{\text{diff},1}(k, n)$, $Y_{\text{diff},2}(k, n)$, ..., $Y_{\text{diff},v}(k, n)$, или при этом промежуточный диффузный сигнал и упомянутые один или более декоррелированных сигналов формируют упомянутые один или более обработанных диффузных сигналов $Y_{\text{diff},1}(k, n)$, $Y_{\text{diff},2}(k, n)$, ..., $Y_{\text{diff},v}(k, n)$.

Например, количество обработанных диффузных сигналов $Y_{\text{diff},1}(k, n)$, $Y_{\text{diff},2}(k, n)$, ..., $Y_{\text{diff},v}(k, n)$ и количество выходных аудиосигналов $Y_1(k, n)$, $Y_2(k, n)$, ..., $Y_v(k, n)$ могут, например, быть равными.

Генерирование упомянутых одного или более декоррелированных сигналов из промежуточного диффузного сигнала может, например, выполняться посредством применения задержек к промежуточному диффузному сигналу, или, например, посредством свертки промежуточного диффузного сигнала с шумовым выбросом, или, например, посредством свертки промежуточного диффузного сигнала с импульсной характеристикой, и т.д. Альтернативно или дополнительно может, например, применяться любой другой известный из уровня техники способ декорреляции.

Для получения v выходных аудиосигналов $Y_1(k, n)$, $Y_2(k, n)$, ..., $Y_v(k, n)$, могут, например, применяться v определений v усилений прямого звука $G_1(k, n)$, $G_2(k, n)$, ..., $G_v(k, n)$ и v применений соответствующего усиления к упомянутым одному или более сигналам прямых компонент $X_{\text{dir}}(k, n)$, чтобы получать v выходных аудиосигналов $Y_1(k, n)$, $Y_2(k, n)$, ..., $Y_v(k, n)$.

Только одиночный сигнал диффузных компонент $X_{\text{diff}}(k, n)$, только одно определение одиночного усиления диффузного звука $Q(k, n)$ и только одно применение усиления диффузного звука $Q(k, n)$ к сигналу диффузных компонент $X_{\text{diff}}(k, n)$ могут, например, требоваться, чтобы получать v выходных аудиосигналов $Y_1(k, n)$, $Y_2(k, n)$, ..., $Y_v(k, n)$. Чтобы достигать декорреляции, способы декорреляции могут применяться только после того, как усиление диффузного звука было уже применено к сигналу диффузных компонент.

Согласно варианту осуществления из фиг. 1а, один и тот же обработанный диффузный сигнал $Y_{\text{diff}}(k, n)$ затем комбинируется с соответствующим одним ($Y_{\text{dir},i}(k, n)$) из обработанных прямых сигналов, чтобы получать соответствующий один ($Y_i(k, n)$) из

выходных аудиосигналов.

Вариант осуществления из фиг. 1а учитывает направление прибытия компонент прямых сигналов упомянутых двух или более входных аудиосигналов $x_1(k, n)$, $x_2(k, n)$, ..., $x_p(k, n)$. Таким образом, выходные аудиосигналы $Y_1(k, n)$, $Y_2(k, n)$, ..., $Y_v(k, n)$ могут генерироваться посредством гибкой регулировки сигналов прямых компонент $X_{dir}(k, n)$ и сигналов диффузных компонент $X_{diff}(k, n)$ в зависимости от направления прибытия. Достигаются усовершенствованные возможности адаптации.

Согласно вариантам осуществления, выходные аудиосигналы $Y_1(k, n)$, $Y_2(k, n)$, ..., $Y_v(k, n)$ могут, например, определяться для каждого время-частотного интервала (k, n) частотно-временной области.

Согласно одному варианту осуществления, модуль 101 генерирования сигнала компонент может, например, быть сконфигурирован с возможностью принимать два или более входных аудиосигналов $x_1(k, n)$, $x_2(k, n)$, ..., $x_p(k, n)$. В другом варианте осуществления, модуль 101 генерирования сигнала компонент может, например, быть сконфигурирован с возможностью принимать три или более входных аудиосигналов $x_1(k, n)$, $x_2(k, n)$, ..., $x_p(k, n)$. Модуль 101 генерирования сигнала компонент может, например, быть сконфигурирован с возможностью разлагать упомянутые два или более (или три или более входных аудиосигналов) $x_1(k, n)$, $x_2(k, n)$, ..., $x_p(k, n)$ на сигнал диффузных компонент $X_{diff}(k, n)$, который не является многоканальным сигналом, и на упомянутые один или более сигналы прямых компонент $X_{dir}(k, n)$. То, что аудиосигнал не является многоканальным сигналом, означает, что сам аудиосигнал не содержит более, чем один аудиоканал. Таким образом, аудиоинформация множества входных аудиосигналов передается внутри упомянутых двух сигналов компонент ($X_{dir}(k, n)$, $X_{diff}(k, n)$) (и возможно в дополнительной вспомогательной информации), что обеспечивает возможность эффективной передачи.

Сигнальный процессор 105, может, например, быть сконфигурирован с возможностью генерировать каждый выходной аудиосигнал $Y_i(k, n)$ из двух или более выходных аудиосигналов $Y_1(k, n)$, $Y_2(k, n)$, ..., $Y_v(k, n)$ посредством определения усиления прямого звука $G_i(k, n)$ для упомянутого выходного аудиосигнала $Y_i(k, n)$, посредством применения упомянутого усиления прямого звука $G_i(k, n)$ к упомянутым одному или более сигналам прямых компонент $X_{dir}(k, n)$, чтобы получать обработанный прямой сигнал $Y_{dir,i}(k, n)$ для упомянутого выходного аудиосигнала $Y_i(k, n)$, и посредством комбинирования упомянутого обработанного прямого сигнала $Y_{dir,i}(k, n)$ для упомянутого выходного аудиосигнала $Y_i(k, n)$ и обработанного диффузного сигнала $Y_{diff}(k, n)$, чтобы генерировать упомянутый выходной аудиосигнал $Y_i(k, n)$. Интерфейс 106 вывода сконфигурирован с возможностью выводить упомянутые два или более выходных аудиосигналов $Y_1(k, n)$, $Y_2(k, n)$, ..., $Y_v(k, n)$. Генерирование двух или более выходных аудиосигналов $Y_1(k, n)$, $Y_2(k, n)$, ..., $Y_v(k, n)$ посредством определения только одиночного обработанного диффузного сигнала $Y_{diff}(k, n)$ является особенно предпочтительным.

Фиг. 1b иллюстрирует устройство для генерирования одного или более выходных аудиосигналов $Y_1(k, n)$, $Y_2(k, n)$, ..., $Y_v(k, n)$ согласно одному варианту осуществления. Устройство осуществляет так называемую сторону "дальнего конца" системы из фиг.

1a.

Устройство из фиг. 1b содержит сигнальный процессор 105, и интерфейс 106 вывода.

Сигнальный процессор 105 сконфигурирован с возможностью принимать сигнал прямых компонент $X_{\text{dir}}(k, n)$, содержащий компоненты прямых сигналов из упомянутых двух или более исходных аудиосигналов $x_1(k, n), x_2(k, n), \dots, x_p(k, n)$ (например, входных аудиосигналов из фиг. 1a). Более того, сигнальный процессор 105 сконфигурирован с возможностью принимать сигнал диффузных компонент $X_{\text{diff}}(k, n)$, содержащий компоненты диффузных сигналов из упомянутых двух или более исходных аудиосигналов $x_1(k, n), x_2(k, n), \dots, x_p(k, n)$. Дополнительно, сигнальный процессор 105 сконфигурирован с возможностью принимать информацию направления, при этом упомянутая информация направления зависит от направления прибытия компонент прямых сигналов упомянутых двух или более входных аудиосигналов.

Сигнальный процессор 105 сконфигурирован с возможностью генерировать один или более обработанных диффузных сигналов $Y_{\text{diff},1}(k, n), Y_{\text{diff},2}(k, n), \dots, Y_{\text{diff},v}(k, n)$ в зависимости от сигнала диффузных компонент $X_{\text{diff}}(k, n)$.

Для каждого выходного аудиосигнала $Y_i(k, n)$ из упомянутых одного или более выходных аудиосигналов $Y_1(k, n), Y_2(k, n), \dots, Y_v(k, n)$, сигнальный процессор 105 сконфигурирован с возможностью определять, в зависимости от направления прибытия, усиление прямого звука $G_i(k, n)$, сигнальный процессор 105 сконфигурирован с возможностью применять упомянутое усиление прямого звука $G_i(k, n)$ к сигналу прямых компонент $X_{\text{dir}}(k, n)$, чтобы получать обработанный прямой сигнал $Y_{\text{dir},i}(k, n)$, и сигнальный процессор 105 сконфигурирован с возможностью комбинировать упомянутый обработанный прямой сигнал $Y_{\text{dir},i}(k, n)$ и один $Y_{\text{diff},i}(k, n)$ из упомянутых одного или более обработанных диффузных сигналов $Y_{\text{diff},1}(k, n), Y_{\text{diff},2}(k, n), \dots, Y_{\text{diff},v}(k, n)$, чтобы генерировать упомянутый выходной аудиосигнал $Y_i(k, n)$.

Интерфейс 106 вывода сконфигурирован с возможностью выводить упомянутые один или более выходных аудиосигналов $Y_1(k, n), Y_2(k, n), \dots, Y_v(k, n)$.

Все конфигурации сигнального процессора 105, описываемые со ссылкой на систему в последующем, также могут осуществляться в устройстве согласно фиг. 1b. Это относится, в частности, к различным конфигурациям модуля 103 модификации сигналов и модуля 104 вычисления функций усиления, которые описываются ниже. То же применяется для различных примеров применений концепций, описанных ниже.

Фиг. 1c иллюстрирует систему согласно другому варианту осуществления. На фиг. 1c, генератор 105 сигналов из фиг. 1a дополнительно содержит модуль 104 вычисления функций усиления для вычисления одной или более функций усиления, при этом каждая функция усиления из упомянутых одной или более функций усиления содержит множество значений аргумента функции усиления, при этом возвращаемое значение функции усиления назначено каждому из упомянутых значений аргумента функции усиления, при этом, когда упомянутая функция усиления принимает одно из упомянутых значений аргумента функции усиления, упомянутая функция усиления сконфигурирована с возможностью возвращать возвращаемое значение функции усиления, которое назначено упомянутому одному из упомянутых значений аргумента функции усиления.

Дополнительно, сигнальный процессор 105 дополнительно содержит модуль 103 модификации сигналов для выбора, в зависимости от направления прибытия, зависящего от направления значения аргумента из значений аргумента функции усиления для

функции усиления из упомянутых одной или более функций усиления, для получения возвращаемого значения функции усиления, которое назначено упомянутому зависящему от направления значению аргумента, от упомянутой функции усиления, и для определения значения усиления, по меньшей мере, одного из упомянутых одного или более выходных аудиосигналов в зависимости от упомянутого возвращаемого значения функции усиления, полученного от упомянутой функции усиления.

Фиг. 1d иллюстрирует систему согласно другому варианту осуществления. На фиг. 1d, генератор 105 сигналов из фиг. 1b дополнительно содержит модуль 104 вычисления функций усиления для вычисления одной или более функций усиления, при этом каждая функция усиления из упомянутых одной или более функций усиления содержит множество значений аргумента функции усиления, при этом возвращаемое значение функции усиления назначено каждому из упомянутых значений аргумента функции усиления, при этом, когда упомянутая функция усиления принимает одно из упомянутых значений аргумента функции усиления, упомянутая функция усиления сконфигурирована с возможностью возвращать возвращаемое значение функции усиления, которое назначено упомянутому одному из упомянутых значений аргумента функции усиления.

Дополнительно, сигнальный процессор 105 дополнительно содержит модуль 103 модификации сигналов для выбора, в зависимости от направления прибытия, зависящего от направления значения аргумента из значений аргумента функции усиления для функции усиления из упомянутых одной или более функций усиления, для получения возвращаемого значения функции усиления, которое назначено упомянутому зависящему от направления значению аргумента, от упомянутой функции усиления, и для определения значения усиления, по меньшей мере, одного из упомянутых одного или более выходных аудиосигналов в зависимости от упомянутого возвращаемого значения функции усиления, полученного от упомянутой функции усиления.

Варианты осуществления обеспечивают запись и воспроизведение пространственного звука таким образом, чтобы акустическое изображение было совместимым с требуемым пространственным изображением, которое определяется, например, посредством видео, которое дополняет аудио на стороне дальнего конца. Некоторые варианты осуществления основываются на записях с помощью массива микрофонов, расположенного в реверберационной стороне ближнего конца. Варианты осуществления обеспечивают, например, акустическое масштабирование, которое является совместимым с визуальным масштабированием камеры. Например, при увеличении масштаба, прямой звук громкоговорителей воспроизводится из направления, где громкоговорители располагались бы в масштабированном визуальном изображении, чтобы визуальное и акустическое изображение были выровнены. Если громкоговорители располагаются вне визуального изображения (или вне требуемой пространственной области) после увеличения масштаба, прямой звук этих громкоговорителей может ослабляться, так как эти громкоговорители более не видны, или, например, так как прямой звук от этих громкоговорителей не требуется. Более того, отношение прямого звука к реверберации может, например, увеличиваться при увеличении масштаба, чтобы имитировать более малый угол раскрытия визуальной камеры.

Варианты осуществления основываются на концепции для разделения записанных сигналов микрофонов на прямой звук источников звука и диффузный звук, например, реверберационный звук, посредством применения двух недавних многоканальных фильтров на стороне ближнего конца. Эти многоканальные фильтры могут, например, основываться на параметрической информации звукового поля, такой как DOA прямого звука. В некоторых вариантах осуществления, разделенные прямой звук и диффузный

звук могут, например, передаваться стороне дальнего конца вместе с параметрической информацией.

Например, на стороне дальнего конца, к извлеченным прямому звуку и диффузному звуку могут, например, применяться конкретные веса, которые регулируют воспроизводимое акустическое изображение, чтобы результирующие выходные аудиосигналы были совместимыми с требуемым пространственным изображением. Эти веса моделируют, например, эффект акустического масштабирования и зависят, например, от направления прибытия (DOA) прямого звука и, например, от коэффициента масштабирования и/или направления просмотра камеры. Окончательные выходные аудиосигналы могут, например, затем получаться посредством суммирования взвешенных прямого звука и диффузного звука.

Обеспеченные концепции реализуют эффективное использование в вышеупомянутом сценарии записи видео с пользовательскими устройствами или в сценарии телеконференцсвязи: Например, в сценарии записи видео, может, например, быть достаточным сохранять или передавать извлеченные прямой звук и диффузный звук (вместо всех сигналов микрофонов), при этом еще имеется возможность управлять воссоздаваемым пространственным изображением.

Это означает, что, если, например, визуальное масштабирование применяется на этапе последующей обработки (цифровое масштабирование), акустическое изображение может все еще модифицироваться соответствующим образом без необходимости сохранять и осуществлять доступ к исходным сигналам микрофонов. В сценарии телеконференцсвязи, предложенные концепции также могут эффективно использоваться, так как извлечение прямого и диффузного звуков может выполняться на стороне ближнего конца, при этом еще имеется возможность управлять воспроизведением пространственного звука (например, изменением установки громкоговорителей) на стороне дальнего конца и выравнивать акустическое и визуальное изображение. Поэтому, только необходимо передавать только несколько аудиосигналов и оцененные направления DOA в качестве вспомогательной информации, при этом вычислительная сложность на стороне дальнего конца является низкой.

Фиг. 2 иллюстрирует систему согласно одному варианту осуществления. Сторона ближнего конца содержит модули 101 и 102. Сторона дальнего конца содержит модуль 105 и 106. Сам модуль 105 содержит модули 103 и 104. Когда ссылка делается на сторону ближнего конца и на сторону дальнего конца, следует понимать, что в некоторых вариантах осуществления, первое устройство может осуществлять сторону ближнего конца (например, содержащую модули 101 и 102), и второе устройство может осуществлять сторону дальнего конца (например, содержащую модули 103 и 104), в то время как в других вариантах осуществления, одиночное устройство осуществляет как сторону ближнего конца, так и сторону дальнего конца, при этом такое одиночное устройство, например, содержит модули 101, 102, 103 и 104.

В частности, фиг. 2 иллюстрирует систему согласно одному варианту осуществления, содержащую модуль 101 генерирования сигнала компонент, модуль 102 оценки параметров, сигнальный процессор 105, и интерфейс 106 вывода. На фиг. 2, сигнальный процессор 105 содержит модуль 104 вычисления функций усиления и модуль 103 модификации сигналов. Сигнальный процессор 105 и интерфейс 106 вывода могут, например, реализовать устройство, как проиллюстрировано посредством фиг. 1b.

На фиг. 2, среди прочего, модуль 102 оценки параметров может, например, быть сконфигурирован с возможностью принимать упомянутые два или более входных аудиосигналов $x_1(k, n)$, $x_2(k, n)$, ... $x_p(k, n)$. Дополнительно модуль 102 оценки параметров

может, например, быть сконфигурирован с возможностью оценивать направление прибытия компонент прямых сигналов упомянутых двух или более входных аудиосигналов $x_1(k, n)$, $x_2(k, n)$, ... $x_p(k, n)$ в зависимости от упомянутых двух или более входных аудиосигналов. Сигнальный процессор 105 может, например, быть

5 сконфигурирован с возможностью принимать информацию направления прибытия, содержащую направление прибытия компонент прямых сигналов упомянутых двух или более входных аудиосигналов, от модуля 102 оценки параметров.

Вход системы из фиг. 2 состоит из M сигналов микрофонов $X_{1...M}(k, n)$ в частотно-временной области (индекс частоты k , временной индекс n). Можно, например,

10 предполагать, что звуковое поле, которое захватывается посредством микрофонов, состоит для каждого (k, n) из плоской волны, распространяющейся в изотропном диффузном поле. Плоская волна моделирует прямой звук источников звука (например, громкоговорителей), в то время как диффузный звук моделирует реверберацию.

Согласно такой модели, сигнал m -ого микрофона может быть записан как

$$15 \quad X_m(k, n) = X_{dir,m}(k, n) + X_{diff,m}(k, n) + X_{n,m}(k, n), \quad (1)$$

где $X_{dir,m}(k, n)$ является измеренным прямым звуком (плоской волной), $X_{diff,m}(k, n)$ является измеренным диффузным звуком, и $X_{n,m}(k, n)$ является компонентой шума (например, собственным шумом микрофона).

20 В модуле 101 генерирования сигнала компонент на фиг. 2 (разложение на прямой/диффузный звук), прямой звук $X_{dir}(k, n)$ и диффузный звук $X_{diff}(k, n)$ извлекается из сигналов микрофонов. Для этой цели могут использоваться, например, информированные многоканальные фильтры, как описано ниже. Для разложения на

25 прямой/диффузный звук, может, например, использоваться конкретная параметрическая информация о звуковом поле, например, DOA прямого звука $\phi(k, n)$. Эта параметрическая информация может, например, оцениваться из сигналов микрофонов в модуле 102 оценки параметров. Помимо DOA $\phi(k, n)$ прямого звука, в некоторых вариантах осуществления, может, например, оцениваться информация расстояния $r(k, n)$. Эта информация расстояния может, например, описывать расстояние между массивом

30 микрофонов и источником звука, который испускает плоскую волну. Для оценки параметров, могут, например, использоваться средства оценки расстояния и/или известные из уровня техники средства оценки DOA. Соответствующие средства оценки могут, например, описываться ниже.

35 Извлеченный прямой звук $X_{dir}(k, n)$, извлеченный диффузный звук $X_{diff}(k, n)$, и оцененная параметрическая информация прямого звука, например, DOA $\phi(k, n)$ и/или расстояние $r(k, n)$, могут, например, затем сохраняться, передаваться стороне дальнего конца, или немедленно использоваться, чтобы генерировать пространственный звук с требуемым пространственным изображением, например, чтобы создавать эффект

40 акустического масштабирования.

Требуемое акустическое изображение, например, эффект акустического масштабирования, генерируется в модуле 103 модификации сигналов с использованием извлеченного прямого звука $X_{dir}(k, n)$, извлеченного диффузного звука $X_{diff}(k, n)$, и оцененной параметрической информации $\phi(k, n)$ и/или $r(k, n)$.

45 Модуль 103 модификации сигналов может, например, вычислять один или более выходных сигналов $Y_j(k, n)$ в частотно-временной области, которые воссоздают акустическое изображение, чтобы оно было совместимым с требуемым пространственным изображением. Например, выходные сигналы $Y_j(k, n)$ имитируют

эффект акустического масштабирования. Эти сигналы могут, в конечном счете, преобразовываться назад во временную область и проигрываться, например, посредством громкоговорителей или наушников. i -ый выходной сигнал $Y_i(k, n)$ вычисляется как взвешенная сумма извлеченного прямого звука $X_{dir}(k, n)$ и диффузного звука $X_{diff}(k, n)$, например,

$$Y_i(k, n) = G_i(k, n)X_{dir}(k, n) + f_i \left\{ \underbrace{QX_{diff}(k, n)}_{Y_{diff}(k, n)} \right\} \quad (2a)$$

$$\equiv Y_{dir,i}(k, n) + Y_{diff,i}(k, n). \quad (2b)$$

В формулах (2a) и (2b), веса $G_i(k, n)$ и Q являются параметрами, которые используются, чтобы создавать требуемое акустическое изображение, например, эффект акустического масштабирования. Например, при увеличении масштаба, параметр Q может уменьшаться, чтобы воспроизводимый диффузный звук ослаблялся.

Более того, с помощью весов $G_i(k, n)$ можно управлять тем, из какого направления прямой звук воспроизводится, чтобы визуальное и акустическое изображение были выровнены. Более того, эффект акустического размытия может выравниваться с прямым звуком.

В некоторых вариантах осуществления, веса $G_i(k, n)$ и Q могут, например, определяться в блоках 201 и 202 выбора усиления. Эти блоки могут, например, выбирать соответствующие веса $G_i(k, n)$ и Q из двух функций усиления, обозначенных посредством g_i и q , в зависимости от оцененной параметрической информации $\varphi(k, n)$ и $r(k, n)$. Выражая математически,

$$G_i(k, n) = g_i(\varphi, r), \quad (3a)$$

$$Q(k, n) = q(r). \quad (3b)$$

В некоторых вариантах осуществления, функции усиления g_i и q могут зависеть от применения и могут, например, генерироваться в модуле 104 вычисления функций усиления. Функции усиления описывают то, какие веса $G_i(k, n)$ и Q должны использоваться в (2a) для заданной параметрической информации $\varphi(k, n)$ и/или $r(k, n)$, чтобы получалось требуемое совместимое пространственное изображение.

Например, при увеличении масштаба с помощью визуальной камеры, функции усиления регулируются, чтобы звук воспроизводился из направлений, где источники являются видимыми на видео. Веса $G_i(k, n)$ и Q и лежащие в основе функции усиления g_i и q дополнительно описываются ниже. Следует отметить, что веса $G_i(k, n)$ и Q и лежащие в основе функции усиления g_i и q могут, например, быть комплекснозначными. Вычисление функций усиления требует информации, такой как коэффициент масштабирования, ширина визуального изображения, требуемое направление просмотра, и установка громкоговорителей.

В других вариантах осуществления, веса $G_i(k, n)$ и Q вычисляются непосредственно внутри модуля 103 модификации сигналов, вместо того, чтобы сначала вычислять функции усиления в модуле 104 и затем выбирать веса $G_i(k, n)$ и Q из вычисленных функций усиления в блоках 201 и 202 выбора усиления.

Согласно вариантам осуществления, более, чем одна плоская волна в расчете на

время-частоту может, например, конкретно обрабатываться. Например, две или более плоские волны в одном и том же частотном диапазоне из двух разных направлений могут, например, прибывать и записываться посредством массива микрофонов в одной и той же точке во времени. Эти две плоские волны могут, каждая, иметь разное
5 направление прибытия. В таких сценариях, компоненты прямых сигналов из упомянутых двух или более плоских волн и их направление прибытий могут, например, учитываться отдельно.

Согласно вариантам осуществления, сигнал прямых компонент $X_{dir1}(k, n)$ и один или более дополнительных сигналов прямых компонент $X_{dir2}(k, n), \dots, X_{dirq}(k, n)$ могут,
10 например, формировать группу из двух или более сигналов прямых компонент $X_{dir1}(k, n), X_{dir2}(k, n), \dots, X_{dirq}(k, n)$, при этом модуль 101 генерирования сигнала компонент может, например, быть сконфигурирован с возможностью генерировать упомянутые один или более дополнительных сигналов прямых компонент $X_{dir2}(k, n), \dots, X_{dirq}(k, n)$,
15 содержащих дополнительные компоненты прямых сигналов упомянутых двух или более входных аудиосигналов $x_1(k, n), x_2(k, n), \dots, x_p(k, n)$.

Направление прибытия и одно или более дополнительных направлений прибытий формируют группу из двух или более направлений прибытий, при этом каждое направление прибытия из группы из упомянутых двух или более направлений прибытий
20 назначено в точности одному сигналу прямых компонент $X_{dir1}(k, n)$ из группы из упомянутых двух или более сигналов прямых компонент $X_{dir1}(k, n), X_{dir2}(k, n), \dots, X_{dirq,m}(k, n)$, при этом количество сигналов прямых компонент из упомянутых двух или более сигналов прямых компонент и количество направлений прибытий из упомянутых двух направлений прибытий является равным.
25

Сигнальный процессор 105 может, например, быть сконфигурирован с возможностью принимать группу из упомянутых двух или более сигналов прямых компонент $X_{dir1}(k, n), X_{dir2}(k, n), \dots, X_{dirq}(k, n)$, и группу из упомянутых двух или более направлений прибытий.

Для каждого выходного аудиосигнала $Y_l(k, n)$ из упомянутых одного или более
30 выходных аудиосигналов $Y_1(k, n), Y_2(k, n), \dots, Y_v(k, n)$,

- Сигнальный процессор 105 может, например, быть сконфигурирован с возможностью определять, для каждого сигнала прямых компонент $X_{dir1}(k, n)$ из группы из упомянутых двух или более сигналов прямых компонент $X_{dir1}(k, n), X_{dir2}(k, n), \dots, X_{dirq}(k, n)$, усиление прямого звука $G_{j,l}(k, n)$ в зависимости от направления прибытия упомянутого сигнала прямых компонент $X_{dir1}(k, n)$,
35

- Сигнальный процессор 105 может, например, быть сконфигурирован с возможностью генерировать группу из двух или более обработанных прямых сигналов $Y_{dir1,l}(k, n), Y_{dir2,l}(k, n), \dots, Y_{dirq,l}(k, n)$ посредством применения, для каждого сигнала прямых компонент $X_{dir1}(k, n)$ из группы из упомянутых двух или более сигналов прямых компонент $X_{dir1}(k, n), X_{dir2}(k, n), \dots, X_{dirq}(k, n)$, усиления прямого звука $G_{j,l}(k, n)$ упомянутого сигнала прямых компонент $X_{dir1}(k, n)$ к упомянутому сигналу прямых компонент $X_{dir1}(k, n)$. И:
40

- Сигнальный процессор 105 может, например, быть сконфигурирован с возможностью комбинировать один $Y_{diff,l}(k, n)$ из упомянутых одного или более обработанных диффузных сигналов $Y_{diff,1}(k, n), Y_{diff,2}(k, n), \dots, Y_{diff,v}(k, n)$ и каждый

обработанный сигнал $Y_{dir,j}(k, n)$ из группы из упомянутых двух или более обработанных сигналов $Y_{dir,1}(k, n), Y_{dir,2}(k, n), \dots, Y_{dir,q}(k, n)$, чтобы генерировать упомянутый выходной аудиосигнал $Y_f(k, n)$.

5 Таким образом, если две или более плоские волны учитываются отдельно, модель формулы (1) становится:

$$X_m(k, n) = X_{dir,1,m}(k, n) + X_{dir,2,m}(k, n) + \dots + X_{dir,q,m}(k, n) + X_{diff,m}(k, n) + X_{n,m}(k, n)$$

и веса могут, например, вычисляться аналогично формулам (2a) и (2b) согласно:

$$10 \begin{aligned} Y_f(k, n) &= G_{1,f}(k, n) X_{dir,1}(k, n) + G_{2,f}(k, n) X_{dir,2}(k, n) + \dots + G_{q,f}(k, n) X_{dir,q}(k, n) + QX_{diff,m}(k, n) \\ &= Y_{dir,1,f}(k, n) + Y_{dir,2,f}(k, n) + \dots + Y_{dir,q,f}(k, n) + Y_{diff,f}(k, n) \end{aligned}$$

Является достаточным, чтобы только малое количество сигналов прямых компонент, сигнал диффузных компонент и вспомогательная информация передавались от стороны ближнего конца стороне дальнего конца. В одном варианте осуществления, количество сигналов (сигнала) прямых компонент из группы из упомянутых двух или более сигналов прямых компонент $X_{dir,1}(k, n), X_{dir,2}(k, n), \dots, X_{dir,q}(k, n)$ плюс 1 меньше, чем количество входных аудиосигналов $x_1(k, n), x_2(k, n), \dots, x_p(k, n)$, которые принимаются интерфейсом 101 приема, (с использованием индексов: $q+1 < p$) "плюс 1" представляет сигнал диффузных компонент $X_{diff}(k, n)$, который необходим.

20 Когда в последующем, обеспечиваются описания в отношении одиночной плоской волны, одиночного направления прибытия и одиночного сигнала прямых компонент, следует понимать, что описанные концепции одинаковым образом применимы к более, чем одной плоской волне, более, чем одному направлению прибытия и более, чем одному сигналу прямых компонент.

25 В последующем, описывается извлечение прямого и диффузного звуков. Обеспечиваются практические реализации модуля 101 генерирования сигнала компонент из фиг. 2, который реализует разложение на прямой/диффузный звук.

В вариантах осуществления, чтобы реализовать совместимое воспроизведение пространственного звука, вывод двух недавно предложенных информированных фильтров с линейно ограниченной минимальной дисперсией (LCMV), описанных в [8] и [9], комбинируются, что обеспечивает возможность точного многоканального извлечения прямого звука и диффузного звука с требуемым произвольной характеристикой при предположении аналогичной модели звукового поля как в DirAC (направленном аудио кодировании). Теперь в последующем описывается конкретный способ комбинирования этих фильтров согласно одному варианту осуществления:

35 Сначала, описывается извлечение прямого звука согласно одному варианту осуществления.

Прямой звук извлекается с использованием недавно предложенного информированного пространственного фильтра, описанного в [8]. Этот фильтр кратко обзревается в последующем и затем определяется таким образом, чтобы он мог использоваться в вариантах осуществления согласно фиг. 2.

45 Оцененный требуемый прямой сигнал $\hat{Y}_{dir,i}(k, n)$ для i -ого канала громкоговорителя в (2b) и фиг. 2 вычисляется посредством применения линейного многоканального фильтра к сигналам микрофонов, например

$$\hat{Y}_{dir,i}(k, n) = \mathbf{w}_{dir,i}^H(k, n) \mathbf{x}(k, n), \quad (4)$$

где вектор $\mathbf{x}(k, n) = [X_1(k, n), \dots, X_M(k, n)]^T$ содержит M сигналов микрофонов и $\mathbf{w}_{dir,i}$

является комплекснозначным весовым вектором. Здесь, веса фильтра минимизируют шумовой и диффузный звук, составляемый микрофонами при захвате прямого звука с требуемым усилением $G_i(k, n)$. Выражая математически, веса, могут, например, вычисляться как

$$\mathbf{w}_{\text{dir},i}(k, n) = \arg \min_{\mathbf{w}} \mathbf{w}^H \Phi_u(k, n) \mathbf{w}, \quad (5)$$

при условии линейного ограничения

$$\mathbf{w}^H \mathbf{a}(k, \varphi) = G_i(k, n). \quad (6)$$

Здесь, $\mathbf{a}(k, \varphi)$ является так называемым вектором распространения массива. m -ый элемент этого вектора является относительной передаточной функцией прямого звука между m -ым микрофоном и опорным микрофоном массива (без потери общности в последующем описании используется первый микрофон в положении \mathbf{d}_1). Этот вектор зависит от DOA $\varphi(k, n)$ прямого звука.

Вектор распространения массива, например, определяется в [8]. В формуле (6) из документа [8], вектор распространения массива определяется согласно

$$\mathbf{a}(k, \varphi_l) = [a_1(k, \varphi_l) \dots a_M(k, \varphi_l)]^T,$$

где φ_l является азимутальным углом направления прибытия l -й плоской волны.

Таким образом, вектор распространения массива зависит от направления прибытия. Если существует или рассматривается только одна плоская волна, индекс l может пропускаться.

Согласно формуле (6) из [8], i -й элемент a_i вектора распространения массива \mathbf{a} описывает фазовый сдвиг l -й плоской волны от первого к i -му микрофону и определяется согласно

$$a_i(k, \varphi_l) = \exp\{j\kappa r_i \sin \varphi_l(k, n)\}.$$

Например, r_i равняется расстоянию между первым и i -м микрофоном, κ обозначает волновое число плоской волны и j является мнимым числом.

Больше информации о векторе распространения массива \mathbf{a} и его элементах a_i может быть найдено в [8], который явно включается сюда по ссылке.

$M \times M$ матрица $\Phi_u(k, n)$ в (5) является матрицей спектральной плотности мощности (PSD) шумового и диффузного звука, которая может определяться, как описано в [8]. Решение для (5) задается посредством

$$\mathbf{w}_{\text{dir},i}(k, n) = \mathbf{h}_{\text{dir}}(k, n) G_i^*(k, n), \quad (7)$$

где

$$\mathbf{h}_{\text{dir}}(k, n) = \Phi_u^{-1}(k, n) \mathbf{a}(k, \varphi) [\mathbf{a}^H(k, \varphi) \Phi_u^{-1} \mathbf{a}(k, \varphi)]^{-1}. \quad (8)$$

Вычисление фильтра требует вектора распространения массива $\mathbf{a}(k, \varphi)$, который может определяться после того, как было оценено DOA $\varphi(k, n)$ прямого звука [8]. Как описано выше, вектор распространения массива и, таким образом, фильтр, зависит от DOA. DOA может оцениваться как описано ниже.

Информированный пространственный фильтр, предложенный в [8], например, извлечение прямого звука с использованием (4) и (7), не может напрямую использоваться в варианте осуществления из фиг. 2. Фактически, для вычисления требуются сигналы микрофонов $\mathbf{x}(k, n)$ также как усиление прямого звука $G_i(k, n)$. Как можно видеть на

фиг. 2, сигналы микрофонов $\mathbf{x}(k, n)$ являются доступными только на стороне ближнего конца, в то время как усиление прямого звука $G_{\lambda}(k, n)$ доступно только на стороне дальнего конца.

Чтобы использовать информированный пространственный фильтр в вариантах осуществления изобретения, обеспечивается модификация, при этом мы подставляем (7) в (4), что ведет к

$$\hat{Y}_{dir,i}(k, n) = G_i(k, n) \hat{X}_{dir}(k, n), \quad (9)$$

где

$$\hat{X}_{dir}(k, n) = \mathbf{h}_{dir}^H(k, n) \mathbf{x}(k, n). \quad (10)$$

Этот модифицированный фильтр $\mathbf{h}_{dir}(k, n)$ является независимым от весов $G_{\lambda}(k, n)$. Таким образом, фильтр может применяться на стороне ближнего конца, чтобы получать прямой звук $\hat{X}_{dir}(k, n)$, который может затем передаваться стороне дальнего конца вместе с оцененными направлениями DOA (и расстоянием) в качестве вспомогательной информации, чтобы обеспечивать полное управление над воспроизведением прямого звука. Прямой звук $\hat{X}_{dir}(k, n)$ может определяться по отношению к опорному микрофону в положении \mathbf{d}_1 . Поэтому, он может также относиться к компонентам прямого звука как $\hat{X}_{dir}(k, n, \mathbf{d}_1)$, и, таким образом:

$$\hat{X}_{dir}(k, n, \mathbf{d}_1) = \mathbf{h}_{dir}^H(k, n) \mathbf{x}(k, n). \quad (10a)$$

Таким образом, согласно одному варианту осуществления, модуль 101 генерирования сигнала компонент может, например, быть сконфигурирован с возможностью генерировать сигнал прямых компонент посредством применения фильтра к упомянутым двум или более входным аудиосигналам согласно

$$\hat{X}_{dir}(k, n) = \mathbf{h}_{dir}^H(k, n) \mathbf{x}(k, n),$$

где k обозначает частоту, и где n обозначает время, где $\hat{X}_{dir}(k, n)$ обозначает сигнал прямых компонент, где $\mathbf{x}(k, n)$ обозначает упомянутые два или более входных аудиосигналов, где $\mathbf{h}_{dir}(k, n)$ обозначает фильтр, где

$$\mathbf{h}_{dir}(k, n) = \Phi_u^{-1}(k, n) \mathbf{a}(k, \varphi) [\mathbf{a}^H(k, \varphi) \Phi_u^{-1} \mathbf{a}(k, \varphi)]^{-1},$$

где $\Phi_u(k, n)$ обозначает матрицу спектральной плотности мощности шумового и диффузного звука упомянутых двух или более входных аудиосигналов, где $\mathbf{a}(k, \varphi)$ обозначает вектор распространения массива, и где φ обозначает азимутальный угол направления прибытия компонент прямых сигналов упомянутых двух или более входных аудиосигналов.

Фиг. 3 иллюстрирует модуль 102 оценки параметров и модуль 101 генерирования сигнала компонент, осуществляющий разложение на прямой/диффузный звук, согласно одному варианту осуществления.

Вариант осуществления, проиллюстрированный посредством фиг. 3, реализует извлечение прямого звука посредством модуля 203 извлечения прямого звука и извлечение диффузного звука посредством модуля 204 извлечения диффузного звука.

Извлечение прямого звука выполняется в модуле 203 извлечения прямого звука

посредством применения весов фильтра к сигналам микрофонов, как задано в (10). Веса фильтра прямого звука вычисляются в блоке 301 вычисления весов прямого звука, который может реализовываться, например, с помощью (8). Усиления $G_f(k, n)$ из, например, уравнения (9), затем применяются на стороне дальнего конца, как показано на фиг. 2.

В последующем, описывается извлечение диффузного звука. Извлечение диффузного звука может, например, осуществляться посредством модуля 204 извлечения диффузного звука из фиг. 3. Веса фильтра диффузного звука вычисляются в блоке 302 вычисления весов диффузного звука из фиг. 3, например, как описано в последующем.

В вариантах осуществления, диффузный звук может, например, извлекаться с использованием пространственного фильтра, который был недавно предложен в [9]. Диффузный звук $X_{diff}(k, n)$ в (2а) и фиг. 2 может, например, оцениваться посредством применения второго пространственного фильтра к сигналам микрофонов, например,

$$\hat{X}_{diff}(k, n) = \mathbf{h}_{diff}^H(k, n)\mathbf{x}(k, n), \quad (11)$$

Чтобы находить оптимальный фильтр для диффузного звука $\mathbf{h}_{diff}(k, n)$, мы рассматриваем недавно предложенный фильтр в [9], который может извлекать диффузный звук с требуемой произвольной характеристикой при минимизации шума на выходе фильтра. Для пространственно белого шума, фильтр задается посредством

$$\mathbf{h}_{diff}(k, n) = \arg \min_{\mathbf{h}} \mathbf{h}^H \mathbf{h} \quad (12)$$

при условии $\mathbf{h}^H \mathbf{a}(k, \varphi) = 0$ и $\mathbf{h}^H \boldsymbol{\gamma}_1(k) = 1$. Первое линейное ограничение обеспечивает, что прямой звук подавляется, в то время как второе ограничение обеспечивает, что в среднем, диффузный звук захватывается с требуемым усилением Q , см. документ [9]. Отметим, что $\boldsymbol{\gamma}_1(k)$ является вектором когерентности диффузного звука, определенным в [9]. Решение для (12) задается посредством

$$\mathbf{h}_{diff}(k, n) = \boldsymbol{\Lambda} \boldsymbol{\gamma}_{diff}(k) [\boldsymbol{\gamma}_{diff}^H(k) \boldsymbol{\Lambda} \boldsymbol{\gamma}_{diff}(k)]^{-1}, \quad (13)$$

где

$$\boldsymbol{\Lambda}(k, \varphi) = \mathbf{I} - \mathbf{a}(k, \varphi) [\mathbf{a}^H(k, \varphi) \mathbf{a}(k, \varphi)]^{-1} \mathbf{a}^H(k, \varphi), \quad (14)$$

где \mathbf{I} является единичной матрицей размера $M \times M$. Фильтр $\mathbf{h}_{diff}(k, n)$ не зависит от весов $G_f(k, n)$ и Q , и, таким образом, он может вычисляться и применяться на стороне ближнего конца, чтобы получать $\hat{X}_{diff}(k, n)$. Таким образом, необходимо передавать только одиночный аудиосигнал стороне дальнего конца, именно $\hat{X}_{diff}(k, n)$, при этом еще имеется возможность полностью управлять воспроизведением пространственного звука диффузного звука.

Фиг. 3, более того, иллюстрирует извлечение диффузного звука согласно одному варианту осуществления. Извлечение диффузного звука выполняется в модуле 204 извлечения диффузного звука посредством применения весов фильтра к сигналам микрофонов, как задано в формуле (11). Веса фильтра вычисляются в блоке 302 вычисления весов диффузного звука, который может реализовываться, например, посредством использования формулы (13).

В последующем, описывается оценка параметров. Оценка параметров может, например, выполняться посредством модуля 102 оценки параметров, в котором может,

например, оцениваться параметрическая информация о записанной звуковой сцене. Эта параметрическая информация используется для вычисления двух пространственных фильтров в модуле 101 генерирования сигнала компонент и для выбора усиления в совместимом пространственном воспроизведении аудио в модуле 103 модификации сигналов.

Сначала, описывается определение/оценка информации DOA.

В последующем описываются варианты осуществления, где модуль (102) оценки параметров содержит средство оценки DOA для прямого звука, например, для плоской волны, которая исходит из положения источника звука и прибывает в массив микрофонов. Без потери общности, предполагается, что для каждого момента времени и частоты существует одиночная плоская волна. Другие варианты осуществления учитывают случаи, когда существует множество плоских волн, и расширение концепций одиночной плоской волны, описанных здесь, на множество плоских волн является непосредственным. Поэтому, настоящее изобретение также охватывает варианты осуществления с множеством плоских волн.

Узкополосные направления DOA могут оцениваться из сигналов микрофонов с использованием одного из известных из уровня техники узкополосных средств оценки DOA, таких как ESPRIT [10] или MUSIC на основе корней [11]. Вместо азимутального угла $\phi(k, n)$, информация DOA также может обеспечиваться в форме пространственной частоты $\mu[k|\phi(k, n)]$, фазового сдвига, или вектора распространения $\mathbf{a}[k|\phi(k, n)]$ для одной или более волн, прибывающих в массив микрофонов. Следует отметить, что информация DOA также может обеспечиваться внешним образом. Например, DOA плоской волны может определяться посредством видеокамеры вместе с алгоритмом распознавания лиц при предположении, что акустическую сцену формируют говорящие люди.

В заключение, следует отметить, что информация DOA также может оцениваться в 3D (в трех измерениях). В этом случае, углы как азимута $\phi(k, n)$, так и возвышения $\vartheta(k, n)$ оцениваются в модуле 102 оценки параметров и DOA плоской волны в таком случае обеспечивается, например, как (ϕ, ϑ) .

Таким образом, когда ссылка делается ниже на азимутальный угол направления DOA, следует понимать, что все описания также применимы к углу возвышения направления DOA, к углу или выведенному из азимутального угла направления DOA, к углу или выведенному из угла возвышения направления DOA или к углу, выведенному из азимутального угла и угла возвышения направления DOA. В более общем, все описания, обеспеченные ниже, равным образом применимы к любому углу, зависящему от DOA.

Теперь, описывается определение/оценка информации расстояния.

Некоторые варианты осуществления относятся к акустическому масштабированию на основе направлений DOA и расстояний. В таких вариантах осуществления, модуль 102 оценки параметров может, например, содержать два подмодуля, например, подмодуль средства оценки DOA, описанный выше, и подмодуль оценки расстояния, который оценивает расстояние от положения записи до источника звука $l(k, n)$. В таких вариантах осуществления, можно, например, предполагать, что каждая плоская волна, которая прибывает в записывающий массив микрофонов, исходит от источника звука и распространяется вдоль прямой линии к массиву (которая также известна как прямой путь распространения).

Существует несколько известных из уровня техники подходов для оценки расстояния с использованием сигналов микрофонов. Например, расстояние до источника может

находиться посредством вычисления отношений мощности между сигналами микрофонов, как описано в [12]. Альтернативно, расстояние до источника $l(k, n)$ в акустических огороженных местах (например, помещениях) может вычисляться на основе оцененного отношения сигнала к диффузному звуку (SDR) [13]. Оценки SDR могут затем комбинироваться с временем реверберации помещения (известным или оцененным с использованием известных из уровня техники способов), чтобы вычислять расстояние. Для высокого SDR, энергия прямого звука является высокой по сравнению с диффузным звуком, что указывает, что расстояние до источника является малым. Когда значение SDR является низким, мощность прямого звука является слабой по сравнению с реверберацией помещения, что указывает большое расстояние до источника.

В других вариантах осуществления, вместо вычисления/оценки расстояния посредством использования модуля вычисления расстояния в модуле 102 оценки параметров, может, например, приниматься внешняя информация расстояния, например, от визуальной системы. Например, могут, например, использоваться известные из уровня техники способы, используемые в системах технического зрения, которые могут обеспечивать информацию расстояния, например, время пролета (ToF), стереоскопическое зрение, и структурированный свет. Например, в камерах ToF, расстояние до источника может вычисляться из измеренного времени пролета сигнала света, испущенного камерой и перемещающегося к источнику и обратно к датчику камеры. Компьютерное стерео зрение, например, использует две точки наблюдения, из которых захватывается визуальное изображение, чтобы вычислять расстояние до источника.

Или, например, могут использоваться камеры структурированного света, где на визуальную сцену проецируется известный шаблон пикселей. Анализ деформаций после проекции обеспечивает возможность визуальной системе оценивать расстояние до источника. Следует отметить, что для совместимого воспроизведения аудиосцены требуется информация расстояния $l(k, n)$ для каждого время-частотного интервала. Если информация расстояния обеспечивается внешним образом посредством визуальной системы, расстояние до источника $l(k, n)$, которое соответствует DOA $\phi(k, n)$, может, например, выбираться как значение расстояния от визуальной системы, которое соответствует этому конкретному направлению $\phi(k, n)$.

В последующем, рассматривается совместимое воспроизведение акустической сцены. Сначала, рассматривается воспроизведение акустической сцены на основе направлений DOA.

Воспроизведение акустической сцены может выполняться таким образом, чтобы оно было совместимым с записанной акустической сценой. Или, воспроизведение акустической сцены может выполняться таким образом, чтобы оно было совместимым с визуальным изображением. Соответствующая визуальная информация может обеспечиваться, чтобы достигать согласованность с визуальным изображением.

Согласованность может, например, достигаться посредством регулировки весов $G_f(k, n)$ и Q в (2a). Согласно вариантам осуществления, модуль 103 модификации сигналов, который может, например, существовать, на стороне ближнего конца, или, как показано на фиг. 2, на стороне дальнего конца, может, например, принимать прямой $\hat{X}_{dir}(k, n)$ и диффузный $\hat{X}_{diff}(k, n)$ звуки в качестве ввода, вместе с оценками DOA $\phi(k, n)$ в качестве вспомогательной информации. На основе этой принятой информации, выходные сигналы $Y_f(k, n)$ для доступной системы воспроизведения могут, например, генерироваться, например, согласно формуле (2a).

В некоторых вариантах осуществления, параметры $G_i(k, n)$ и Q выбираются в блоках 201 и 202 выбора усиления, соответственно, из двух функций усиления $g_i(\phi(k, n))$ и $q(k, n)$, обеспеченных посредством модуля 104 вычисления функций усиления.

5 Согласно одному варианту осуществления, $G_i(k, n)$ может, например, выбираться на основе только информации DOA и Q может, например, иметь постоянное значение. В других вариантах осуществления, однако, другой вес $G_i(k, n)$ может, например, определяться на основе дополнительной информации, и вес Q может, например, определяться переменным образом.

10 Сначала, рассматриваются варианты осуществления, которые реализуют согласованность с записанной акустической сценой. После этого, рассматриваются варианты осуществления, которые реализуют согласованность с информацией изображения/с визуальным изображением.

15 В последующем, описывается вычисление весов $G_i(k, n)$ и Q , чтобы воспроизводить акустическую сцену, которая является совместимой с записанной акустической сценой, например, таким образом, чтобы слушатель, расположенный в зоне наилучшего восприятия системы воспроизведения, воспринимал источники звука как прибывающие из направлений DOA источников звука в записанной звуковой сцене, имеющими такую же мощность как в записанной сцене, и воспроизводящими такое же восприятие

20 объемного диффузного звука. Для известной установки громкоговорителей, воспроизведение источника звука из направления $\phi(k, n)$ может, например, достигаться посредством выбора усиления прямого звука $G_i(k, n)$ в блоке 201 выбора усиления ("выбор усиления прямого звука") из фиксированной таблицы поиска, обеспеченной модулем 104 вычисления функций

25 усиления для оцененного DOA $\phi(k, n)$, что может быть записано как

$$G_i(k, n) = g_i(\phi(k, n)), \quad (15)$$

где $g_i(\phi) = p_i(\phi)$ является функцией, возвращающей усиление панорамирования по всем направлениям DOA для i -ого громкоговорителя. Функция усиления

30 панорамирования $p_i(\phi)$ зависит от установки громкоговорителей и схемы панорамирования.

Пример функции усиления панорамирования $p_i(\phi)$, как определено векторным амплитудным панорамированием (VBAP) [14], для левого и правого громкоговорителя

35 в стерео воспроизведении показан на фиг. 5(a).

На фиг. 5(a), проиллюстрирован пример функции усиления панорамирования VBAP $p_{b,i}$ для стерео установки, и на фиг. 5(b) и иллюстрируются усиления панорамирования для совместимого воспроизведения.

40 Например, если прямой звук прибывает из $\phi(k, n) = 30^\circ$, усиление правого громкоговорителя равняется $G_i(k, n) = g_i(30^\circ) = p_i(30^\circ) = 1$ и усиление левого громкоговорителя равняется $G_i(k, n) = g_i(30^\circ) = p_i(30^\circ) = 0$. Для прямого звука, прибывающего из $\phi(k, n) = 0^\circ$, окончательные усиления стерео громкоговорителя равняются $G_i(k, n) = G_i(k, n) = \sqrt{0.5}$.

45 В одном варианте осуществления, функция усиления панорамирования, например, $p_i(\phi)$, может, например, быть передаточной функцией головы (HRTF) в случае бинаурального воспроизведения звука.

Например, если HRTF $g_i(\phi) = p_i(\phi)$ возвращает комплексные значения, то усиление

прямого звука $G_i(k, n)$, выбранное в блоке 201 выбора усиления, может, например, быть комплекснозначным.

Если должны генерироваться три или более выходных аудиосигналов, могут, например, использоваться соответствующие известные из уровня техники концепции панорамирования, чтобы панорамировать входной сигнал на три или более выходных аудиосигналов. Например, может использоваться VBAR для трех или более выходных аудиосигналов.

В совместимом воспроизведении акустической сцены, мощность диффузного звука должна оставаться такой же как в записанной сцене. Поэтому, для системы громкоговорителей, например, с равноотстоящими громкоговорителями, усиление диффузного звука имеет постоянное значение:

$$Q = q_i = \frac{1}{\sqrt{I}} \quad (16)$$

где I является количеством выходных каналов громкоговорителей. Это означает, что модуль 104 вычисления функций усиления обеспечивает одиночное выходное значение для i -ого громкоговорителя (или канала наушника) в зависимости от количества громкоговорителей, доступных для воспроизведения, и эти значения используются в качестве усиления диффузного звука Q по всем частотам. Окончательный диффузный звук $Y_{diff,i}(k, n)$ для i -ого канала громкоговорителя получается посредством декоррелирования $Y_{diff}(k, n)$, полученного в (2b).

Таким образом, воспроизведение акустической сцены, которое является совместимым с записанной акустической сценой, может достигаться, например, посредством определения усиления для каждого из выходных аудиосигналов в зависимости, например, от направления прибытия, посредством применения множества определенных усиления $G_i(k, n)$ к сигналу прямого звука $\hat{X}_{dir}(k, n)$, чтобы определять множество прямых компонент выходного сигнала $\hat{Y}_{dir,i}(k, n)$, посредством применения определенного усиления Q к сигналу диффузного звука $\hat{X}_{diff}(k, n)$, чтобы получать диффузную компоненту выходного сигнала $\hat{Y}_{diff}(k, n)$, и посредством комбинирования каждой из множества прямых компонент выходного сигнала $\hat{Y}_{dir,i}(k, n)$ с диффузной компонентой выходного сигнала $\hat{Y}_{diff}(k, n)$, чтобы получать упомянутые один или более выходных аудиосигналов $Y_i(k, n)$.

Теперь, описывается генерирование выходных аудиосигналов согласно вариантам осуществления, которое достигает согласованность с визуальной сценой. В частности, описывается вычисление весов $G_i(k, n)$ и Q согласно вариантам осуществления, которые используются, чтобы воспроизводить акустическую сцену, которая является совместимой с визуальной сценой. Целью является воссоздавать акустическое изображение, в котором прямой звук от источника воспроизводится из направления, где источник является видимым на видео/изображении.

Может рассматриваться геометрия, как изображено на фиг. 4, где I соответствует направлению просмотра визуальной камеры. Без потери общности, мы можем определять I на оси u системы координат.

Азимут направления DOA прямого звука в изображенной системе координат (x, y)

задается посредством $\varphi(k, n)$ и местоположение источника на оси x задается посредством $x_g(k, n)$. Здесь, предполагается, что все источники звука располагаются на одном и том же расстоянии g до оси x , например, положения источников располагаются на левой пунктирной линии, которая упоминается в оптике как фокальная плоскость. Следует
 5 отметить, что это предположение делается, только чтобы обеспечивать, чтобы визуальное и акустическое изображения были выровненными, и фактическое значение расстояния g не необходимо для представленной обработки.

На стороне воспроизведения (стороне дальнего конца), устройство отображения располагается в b и положение источника на устройстве отображения задается
 10 посредством $x_b(k, n)$. Более того, x_d является размером устройства отображения (или, в некоторых вариантах осуществления, например, x_d обозначает половину размера устройства отображения), φ_d является, соответствующим максимальным визуальным углом, S является зоной наилучшего восприятия системы воспроизведения звука, и
 15 $\varphi_b(k, n)$ является углом, из которого прямой звук должен воспроизводиться, чтобы визуальное и акустическое изображения были выровнены. $\varphi_b(k, n)$ зависит от $x_b(k, n)$ и от расстояния между зоной наилучшего восприятия S и устройством отображения, расположенным в b . Более того, $x_b(k, n)$ зависит от нескольких параметров, таких как
 20 расстояние g источника от камеры, размер датчика изображений, и размер устройства отображения x_d . К сожалению, по меньшей мере, некоторые из этих параметров часто являются неизвестными на практике, так что $x_b(k, n)$ и $\varphi_b(k, n)$ не могут определяться для заданного DOA $\varphi_g(k, n)$. Однако при предположении, что оптическая система
 является линейной, согласно формуле (17):

$$\tan \varphi_b(k, n) = c \tan \varphi(k, n), \quad (17)$$

где c является неизвестной постоянной, компенсирующей вышеупомянутые неизвестные параметры. Следует отметить, что c является постоянной, только если все
 положения источников имеют одно и то же расстояние g до оси x .

В последующем, предполагается, что c является параметром калибровки, который
 30 должен регулироваться во время этапа калибровки до тех пор, когда визуальное и акустическое изображения будут совместимыми. Чтобы выполнять калибровку, источники звука должны располагаться на фокальной плоскости и значение c находится таким образом, чтобы визуальное и акустическое изображения были выровнены. Когда
 35 откалибровано, значение c остается неизменным и угол, из которого прямой звук должен воспроизводиться, задается посредством

$$\varphi_b(k, n) = \tan^{-1}\{c \tan \varphi(k, n)\}. \quad (18)$$

Чтобы обеспечивать, что как акустическая, так и визуальная сцены являются
 40 совместимыми, исходная функция панорамирования $p_i(\varphi)$ модифицируется в совместимую (модифицированную) функцию панорамирования $p_{b,i}(\varphi)$. Усиление прямого звука $G_i(k, n)$ теперь выбирается согласно

$$G_i(k, n) = g_i(\varphi(k, n)), \quad (19)$$

$$g_i(\varphi) = p_{b,i}(\varphi), \quad (20)$$

где $p_{b,i}(\varphi)$ является совместимой функцией панорамирования, возвращающей усиления панорамирования для i -ого громкоговорителя по всем возможным направлениям DOA

источников. Для фиксированного значения c , такая совместимая функция панорамирования вычисляется в модуле 104 вычисления функций усиления из исходной (например, VBAR) таблицы усиления панорамирования как

$$p_{b,i}(\varphi) = p_i(\tan^{-1}\{c \tan \varphi\}). \quad (21)$$

Таким образом, в вариантах осуществления, сигнальный процессор 105 может, например, быть сконфигурирован с возможностью определять, для каждого выходного аудиосигнала из упомянутых одного или более выходных аудиосигналов, таким образом, что усиление прямого звука $G_i(k, n)$ определяется согласно

$$G_i(k, n) = p_i(\tan^{-1}\{c \tan(\phi(k, n))\}),$$

где i обозначает индекс упомянутого выходного аудиосигнала, где k обозначает частоту, и где n обозначает время, где $G_i(k, n)$ обозначает усиление прямого звука, где $\phi(k, n)$ обозначает угол, зависящий от направления прибытия (например, азимутальный угол направления прибытия), где c обозначает постоянное значение, и где p_i обозначает функцию панорамирования.

В вариантах осуществления, усиление прямого звука $G_i(k, n)$ выбирается в блоке 201 выбора усиления на основе оцененного DOA $\phi(k, n)$ из фиксированной таблицы поиска, обеспеченной модулем 104 вычисления функций усиления, которая вычисляется один раз (после этапа калибровки) с использованием (19).

Таким образом, согласно одному варианту осуществления, сигнальный процессор 105 может, например, быть сконфигурирован с возможностью получать, для каждого выходного аудиосигнала из упомянутых одного или более выходных аудиосигналов, усиление прямого звука для упомянутого выходного аудиосигнала из таблицы поиска в зависимости от направления прибытия.

В одном варианте осуществления, сигнальный процессор 105 вычисляет таблицу поиска для функции усиления прямого звука $g_i(k, n)$. Например, для каждой возможной полной степени, например, $1^\circ, 2^\circ, 3^\circ, \dots$, для значения азимута ϕ направления DOA, усиление прямого звука $G_i(k, n)$ может вычисляться и сохраняться заранее. Затем, когда принимается текущее значение азимута ϕ направления прибытия, сигнальный процессор 105 считывает усиление прямого звука $G_i(k, n)$ для текущего значения азимута ϕ из таблицы поиска. (Текущее значение азимута ϕ , может, например, быть значением аргумента таблицы поиска; и усиление прямого звука $G_i(k, n)$ может, например, быть возвращаемым значением таблицы поиска). Вместо азимута ϕ направления DOA, в других вариантах осуществления, таблица поиска может вычисляться для любого угла, зависящего от направления прибытия. Это имеет преимущество в том, что значение усиления не всегда должно вычисляться для каждой точки во времени, или для каждого время-частотного интервала, но вместо этого, таблица поиска вычисляется однократно и затем, для принятого угла ϕ , усиление прямого звука $G_i(k, n)$ считывается из таблицы поиска.

Таким образом, согласно одному варианту осуществления, сигнальный процессор 105 может, например, быть сконфигурирован с возможностью вычислять таблицу поиска, при этом таблица поиска содержит множество записей, при этом каждая из записей содержит значение аргумента таблицы поиска и возвращаемое значение таблицы поиска, которое назначено упомянутому значению аргумента. Сигнальный процессор 105 может, например, быть сконфигурирован с возможностью получать одно из возвращаемых значений таблицы поиска из таблицы поиска посредством выбора одного

из значений аргумента таблицы поиска для таблицы поиска в зависимости от направления прибытия. Дополнительно, сигнальный процессор 105 может, например, быть сконфигурирован с возможностью определять значение усиления для, по меньшей мере, одного из упомянутых одного или более выходных аудиосигналов в зависимости от упомянутого одного из возвращаемых значений таблицы поиска, полученного из таблицы поиска.

Сигнальный процессор 105 может, например, быть сконфигурирован с возможностью получать другое одно из возвращаемых значений таблицы поиска из (той же) таблицы поиска посредством выбора другого одного из значений аргумента таблицы поиска в зависимости от другого направления прибытия, чтобы определять другое значение усиления. Например, сигнальный процессор может, например, принимать дополнительную информацию направления, например, в более поздний момент во времени, которая зависит от упомянутого дополнительного направления прибытия.

Пример функций панорамирования VBAR и совместимых функций усиления панорамирования показан на фиг. 5(a) и 5(b).

Следует отметить, что вместо пересчета таблиц усиления панорамирования, можно альтернативно вычислять DOA $\varphi_b(k, n)$ для устройства отображения и применять его в исходной функции панорамирования как $\varphi_A(\varphi_b(k, n))$. Это является верным, так как имеет место следующее отношение:

$$p_{b,i}(\varphi(k, n)) \equiv p_i(\varphi_b(k, n)). \quad (22)$$

Однако это требует, чтобы модуль 104 вычисления функций усиления также принимал оцененные направления DOA $\varphi(k, n)$ в качестве ввода, и пересчет DOA, например, выполняемый согласно формуле (18), будет затем выполняться для каждого временного индекса n .

Что касается воспроизведения диффузного звука, акустические и визуальные изображения совместимо воссоздаются, когда обрабатываются таким же образом, как описано для случая без визуальных изображений, например, когда мощность диффузного звука остается такой же, что и мощность диффузного звука в записанной сцене и сигналы громкоговорителей являются некоррелированными версиями $Y_{diff,i}(k, n)$. Для равноотстоящих громкоговорителей, усиление диффузного звука имеет постоянное значение, например, заданное посредством формулы (16). Как результат, модуль 104 вычисления функций усиления обеспечивает одиночное выходное значение для i -ого громкоговорителя (или канала наушника), которое используется в качестве усиления диффузного звука Q по всем частотам. Окончательный диффузный звук $Y_{diff,i}(k, n)$ для i -ого канала громкоговорителя получается посредством декоррелирования $Y_{diff,i}(k, n)$, например, как задается посредством формулы (2b).

Теперь, рассматриваются варианты осуществления, где обеспечивается акустическое масштабирование на основе направлений DOA. В таких вариантах осуществления, может рассматриваться обработка для акустического масштабирования, которая является совместимой с визуальным масштабированием. Это совместимое аудиовизуальное масштабирование достигается посредством регулировки весов $G_i(k, n)$ и Q , например, используемых в формуле (2a), как изображено в модуле 103 модификации сигналов из фиг. 2.

В одном варианте осуществления, усиление прямого звука $G_i(k, n)$ может, например, выбираться в блоке 201 выбора усиления от функции усиления прямого звука $g(k, n)$, вычисленной в модуле 104 вычисления функций усиления, на основе направлений DOA,

оцененных в модуле 102 оценки параметров. Усиление диффузного звука Q выбирается в блоке 202 выбора усиления от функции усиления диффузного звука $q(\beta)$, вычисленной в модуле 104 вычисления функций усиления. В других вариантах осуществления, усиление прямого звука $G_f(k, n)$ и усиление диффузного звука Q вычисляются модулем 103

5 модификации сигналов без вычисления сначала соответствующих функций усиления и затем выбора усилений.

Следует отметить, что в отличие от вышеописанного варианта осуществления, функция усиления диффузного звука $q(\beta)$ определяется на основе коэффициента масштабирования β . В вариантах осуществления, информация расстояния не
10 используется, и, таким образом, в таких вариантах осуществления, она не оценивается в модуле 102 оценки параметров.

Чтобы выводить параметры масштабирования $G_f(k, n)$ и Q в (2а), учитывается геометрия из фиг. 4. Параметры, обозначенные на фигуре, являются аналогичными параметрам, описанным по отношению к фиг. 4 в варианте осуществления выше.

15 Аналогично вышеописанному варианту осуществления, предполагается, что все источники звука располагаются на фокальной плоскости, которая располагается параллельно оси x на расстоянии g . Следует отметить, что некоторые системы с автофокусировкой являются способными обеспечивать g , например, расстояние до фокальной плоскости. Это обеспечивает возможность предполагать, что все источники
20 в изображении являются отчетливыми. На стороне воспроизведения (дальнего конца), DOA $\varphi_b(k, n)$ и положение $x_b(k, n)$ на устройстве отображения зависят от многих параметров, таких как расстояние g источника от камеры, размер датчика изображений, размер устройства отображения x_c и коэффициент масштабирования камеры (например,
25 угол раскрытия камеры) β . При предположении, что оптическая система является линейной, согласно формуле (23):

$$\tan \varphi_b(k, n) = \beta c \tan \varphi(k, n), \quad (23)$$

где c является параметром калибровки, компенсирующим неизвестные оптические
30 параметры и $\beta \geq 1$ является управляемым пользователем коэффициентом масштабирования. Следует отметить, что в визуальной камере, увеличение масштаба с коэффициентом β является эквивалентным умножению $x_b(k, n)$ на β . Более того, c является постоянной, только если все положения источников имеют одно и то же расстояние g до оси x . В этом случае, c может рассматриваться как параметр калибровки,
35 который регулируется однократно таким образом, чтобы визуальное и акустическое изображения были выровнены. Усиление прямого звука $G_f(k, n)$ выбирается от функции усиления прямого звука $g_f(\varphi)$ как

$$G_i(k, n) = g_i(\varphi(k, n)), \quad (24)$$

$$40 \quad g_i(\varphi) = p_{b,i}(\varphi) w_b(\varphi), \quad (25)$$

где $p_{b,i}(\varphi)$ обозначает функцию усиления панорамирования и $w_b(\varphi)$ является оконной функцией усиления для совместимого аудиовизуального масштабирования. Функция усиления панорамирования для совместимого аудиовизуального масштабирования
45 вычисляется в модуле 104 вычисления функций усиления из исходной (например, VBAR) функции усиления панорамирования $p_f(\varphi)$ как

$$p_{b,i}(\varphi) = p_i(\tan^{-1}[\beta c \tan \varphi]), \quad (26)$$

Таким образом, усиление прямого звука $G_i(k, n)$, например, выбираемое в блоке 201 выбора усиления, определяется на основе оцененного DOA $\phi(k, n)$ из таблицы поиска для панорамирования, вычисленной в модуле 104 вычисления функций усиления, которая является фиксированной, если β не изменяется. Следует отметить, что, в некоторых вариантах осуществления, $p_{b,i}(\phi)$ должна пересчитываться, например, посредством использования формулы (26) каждый раз, когда коэффициент масштабирования β модифицируется.

Иллюстративные стерео функции усиления панорамирования для $\beta=1$ и $\beta=3$ показаны на фиг. 6 (см. фиг. 6(a) и фиг. 6(b)). В частности, фиг. 6(a) иллюстрирует иллюстративную функцию усиления панорамирования $p_{b,i}$ для $\beta=1$; фиг. 6(b) иллюстрирует усиления панорамирования после масштабирования с $\beta=3$; и фиг. 6(c) иллюстрирует усиления панорамирования после масштабирования с $\beta=3$ с угловым сдвигом.

Как можно видеть в примере, когда прямой звук прибывает из $\phi(k, n)=10^\circ$, усиление панорамирования для левого громкоговорителя увеличивается для больших значений β , в то время как функция панорамирования для правого громкоговорителя и $\beta=3$ возвращает более малое значение, чем для $\beta=1$. Такое панорамирование эффективно перемещает воспринимаемое положение источника больше к внешним направлениям, когда коэффициент масштабирования β увеличивается.

Согласно вариантам осуществления, сигнальный процессор 105 может, например, быть сконфигурирован с возможностью определять два или более выходных аудиосигналов. Для каждого выходного аудиосигнала из упомянутых двух или более выходных аудиосигналов, функция усиления панорамирования назначена упомянутому выходному аудиосигналу.

Функция усиления панорамирования каждого из упомянутых двух или более выходных аудиосигналов содержит множество значений аргумента функции панорамирования, при этом возвращаемое значение функции панорамирования назначено каждому из упомянутых значений аргумента функции панорамирования, при этом, когда упомянутая функция панорамирования принимает одно из упомянутых значений аргумента функции панорамирования, упомянутая функция панорамирования сконфигурирована с возможностью возвращать возвращаемое значение функции панорамирования, которое назначено упомянутому одному из упомянутых значений аргумента функции панорамирования. и

Сигнальный процессор 105 сконфигурирован с возможностью определять каждый из упомянутых двух или более выходных аудиосигналов в зависимости от зависящего от направления значения аргумента из значений аргумента функции панорамирования для функции усиления панорамирования, которая назначена упомянутому выходному аудиосигналу, при этом упомянутое зависящее от направления значение аргумента зависит от направления прибытия.

Согласно одному варианту осуществления, функция усиления панорамирования каждого из упомянутых двух или более выходных аудиосигналов имеет один или более глобальных максимумов, являющихся одним из значений аргумента функции панорамирования, при этом для каждого из упомянутых одного или более глобальных максимумов каждой функции усиления панорамирования, не существует никакое другое значение аргумента функции панорамирования, для которого упомянутая функция усиления панорамирования возвращает более большое возвращаемое значение функции панорамирования, чем для упомянутых глобальных максимумов.

Для каждой пары из первого выходного аудиосигнала и второго выходного

аудиосигнала из упомянутых двух или более выходных аудиосигналов, по меньшей мере, один из упомянутых одного или более глобальных максимумов функции усиления панорамирования первого выходного аудиосигнала отличается от любого из упомянутых одного или более глобальных максимумов функции усиления панорамирования второго выходного аудиосигнала.

Формулируя кратко, функции панорамирования осуществляются таким образом, чтобы (по меньшей мере, один из) глобальные максимумы разных функций панорамирования отличались.

Например, на фиг. 6(a), локальные максимумы для $p_{b,l}(\varphi)$ находятся в диапазоне от -45° до -28° и локальные максимумы для $p_{b,r}(\varphi)$ находятся в диапазоне от $+28^\circ$ до $+45^\circ$ и, таким образом, глобальные максимумы отличаются.

Например, на фиг. 6(b), локальные максимумы для $p_{b,l}(\varphi)$ находятся в диапазоне от -45° до -8° и локальные максимумы для $p_{b,r}(\varphi)$ находятся в диапазоне от $+8^\circ$ до $+45^\circ$ и, таким образом, глобальные максимумы также отличаются.

Например, на фиг. 6(c), локальные максимумы для $p_{b,l}(\varphi)$ находятся в диапазоне от -45° до $+2^\circ$ и локальные максимумы для $p_{b,r}(\varphi)$ находятся в диапазоне от $+18^\circ$ до $+45^\circ$ и, таким образом, глобальные максимумы также отличаются.

Функция усиления панорамирования может, например, осуществляться как таблица поиска.

В таком варианте осуществления, сигнальный процессор 105 может, например, быть сконфигурирован с возможностью вычислять таблицу поиска панорамирования для функции усиления панорамирования, по меньшей мере, одного из выходных аудиосигналов.

Таблица поиска панорамирования каждого выходного аудиосигнала из упомянутых, по меньшей мере, одного из выходных аудиосигналов может, например, содержать множество записей, при этом каждая из записей содержит значение аргумента функции панорамирования для функции усиления панорамирования упомянутого выходного аудиосигнала и возвращаемое значение функции панорамирования для функции усиления панорамирования, которое назначено упомянутому значению аргумента функции панорамирования, при этом сигнальный процессор 105 сконфигурирован с возможностью получать одно из возвращаемых значений функции панорамирования из упомянутой таблицы поиска панорамирования посредством выбора, в зависимости от направления прибытия, зависящего от направления значения аргумента из таблицы поиска панорамирования, и при этом сигнальный процессор 105 сконфигурирован с возможностью определять значение усиления для упомянутого выходного аудиосигнала в зависимости от упомянутого одного из возвращаемых значений функции панорамирования, полученного из упомянутой таблицы поиска панорамирования.

В последующем, описываются варианты осуществления, которые используют окно прямого звука. Согласно таким вариантам осуществления, окно прямого звука для совместимого масштабирования $w_b(\varphi)$ вычисляется согласно

$$w_b(\varphi) = w(\tan^{-1}[\beta c \tan \varphi]), \quad (27)$$

где $w_b(\varphi)$ является оконной функцией усиления для акустического масштабирования, которая ослабляет прямой звук, если источник отображается в положение вне визуального изображения для коэффициента масштабирования β .

Оконная функция $w(\varphi)$ может, например, устанавливаться для $\beta=1$, чтобы прямой звук источников, которые находятся вне визуального изображения, уменьшался на

требуемый уровень, и она может пересчитываться, например, посредством использования формулы (27), каждый раз, когда параметр масштабирования изменяется. Следует отметить, что $w_b(\varphi)$ является одной и той же для всех каналов

громкоговорителей. Иллюстративные оконные функции для $\beta=1$ и $\beta=3$ показаны на

5 фиг. 7(a-b), где для увеличенного значения β ширина окна уменьшается.

На фиг. 7 проиллюстрированы примеры совместимых оконных функций усиления. В частности, фиг. 7(a) иллюстрирует оконную функцию усиления w_b без

масштабирования (коэффициент масштабирования $\beta=1$), фиг. 7(b) иллюстрирует оконную функцию усиления после масштабирования (коэффициент масштабирования $\beta=3$), фиг. 10 7(c) иллюстрирует оконную функцию усиления после масштабирования (коэффициент масштабирования $\beta=3$) с угловым сдвигом. Например, угловой сдвиг может реализовать вращение окна в направлении просмотра.

Например, на фиг. 7(a), 7(b) и 7(c) оконная функция усиления возвращает усиление, равное 1, если DOA φ располагается внутри окна, оконная функция усиления возвращает 15 усиление, равное 0.18, если φ располагается вне окна, и оконная функция усиления возвращает усиление между 0.18 и 1, если φ располагается на границе окна.

Согласно вариантам осуществления, сигнальный процессор 105 сконфигурирован с возможностью генерировать каждый выходной аудиосигнал из упомянутых одного или более выходных аудиосигналов в зависимости от оконной функции усиления.

20 Оконная функция усиления сконфигурирована с возможностью возвращать возвращаемое значение оконной функции при приеме значения аргумента оконной функции.

Если значение аргумента оконной функции больше, чем нижний порог окна и меньше, чем верхний порог окна, оконная функция усиления сконфигурирована с возможностью 25 возвращать возвращаемое значение оконной функции, которое больше, чем любое возвращаемое значение оконной функции, возвращаемое оконной функцией усиления, если значение аргумента оконной функции меньше, чем нижний порог, или больше, чем верхний порог.

Например, в формуле (27)

$$30 \quad w_b(\varphi) = w(\tan^{-1}[\beta c \tan \varphi]),$$

азимутальный угол направления прибытия φ является значением аргумента оконной функции для оконной функции усиления $w_b(\varphi)$. Оконная функция усиления $w_b(\varphi)$ 35 зависит от информации масштабирования, здесь, коэффициента масштабирования β .

Чтобы объяснить определение оконной функции усиления, можно сослаться на фиг. 7(a).

Если азимутальный угол направления DOA φ больше, чем -20° (нижний порог) и меньше, чем $+20^\circ$ (верхний порог), все значения, возвращаемые оконной функцией 40 усиления, больше, чем 0.6. В противном случае, если азимутальный угол направления DOA φ меньше, чем -20° (нижний порог) или больше, чем $+20^\circ$ (верхний порог), все значения, возвращаемые оконной функцией усиления, меньше, чем 0.6.

В одном варианте осуществления, сигнальный процессор 105 сконфигурирован с возможностью принимать информацию масштабирования. Более того сигнальный процессор 105 сконфигурирован с возможностью генерировать каждый выходной аудиосигнал из упомянутых одного или более выходных аудиосигналов в зависимости от оконной функции усиления, при этом оконная функция усиления зависит от информации масштабирования.

Это можно видеть для (модифицированных) оконных функций усиления из фиг. 7(b)

и фиг. 7(с), если другие значения рассматриваются как нижний/верхний пороги или если другие значения рассматриваются как возвращаемые значения. На фиг. 7(а), 7(б) и 7(с), можно видеть, что оконная функция усиления зависит от информации масштабирования: коэффициента масштабирования β .

5 Оконная функция усиления может, например, осуществляться как таблица поиска. В таком варианте осуществления, сигнальный процессор 105 сконфигурирован с возможностью вычислять таблицу поиска окна, при этом таблица поиска окна содержит множество записей, при этом каждая из записей содержит значение аргумента оконной функции для оконной функции усиления и возвращаемое значение оконной функции
10 для оконной функции усиления, которое назначено упомянутому значению аргумента оконной функции. Сигнальный процессор 105 сконфигурирован с возможностью получать одно из возвращаемых значений оконной функции из таблицы поиска окна посредством выбора одного из значений аргумента оконной функции для таблицы поиска окна в зависимости от направления прибытия. Более того, сигнальный процессор
15 105 сконфигурирован с возможностью определять значение усиления для, по меньшей мере, одного из упомянутых одного или более выходных аудиосигналов в зависимости от упомянутого одного из возвращаемых значений оконной функции, полученного из таблицы поиска окна.

В дополнение к концепции масштабирования, функции окна и панорамирования
20 могут сдвигаться на угол сдвига θ . Этот угол может соответствовать либо вращению направления просмотра камеры 1, либо перемещению внутри визуального изображения по аналогии с цифровым масштабированием в камерах. В первом случае, угол вращения камеры пересчитывается для угла на устройстве отображения, например, аналогично формуле (23). Во втором случае, θ может быть прямым сдвигом функций окна и
25 панорамирования (например, $w_b(\phi)$ и $p_{b,l}(\phi)$) для совместимого акустического масштабирования. Иллюстративный пример сдвига обеих функций изображен на фиг. 5(с) и 6(с).

Следует отметить, что вместо пересчета функций усиления панорамирования и окна, можно вычислять DOA $\phi_b(k, n)$ для устройства отображения, например, согласно
30 формуле (23), и применять его в исходных функциях панорамирования и окна как $p_l(\phi)$ и $w(\phi_b)$, соответственно. Такая обработка является эквивалентной, так как имеют место следующие отношения:

$$35 \quad p_{b,i}(\varphi(k, n)) \equiv p_i(\varphi_b(k, n)), \quad (28)$$

$$w_b(\varphi(k, n)) \equiv w(\varphi_b(k, n)). \quad (29)$$

Однако это требует, чтобы модуль 104 вычисления функций усиления принимал оцененные направления DOA $\phi(k, n)$ в качестве ввода, и пересчет DOA, например, согласно формуле (18), может, например, выполняться в каждом последовательном
40 временном кадре, независимо, изменилось ли β или нет.

Что касается диффузного звука, вычисление функции усиления диффузного звука $q(\beta)$, например, в модуле 104 вычисления функций усиления, требует только знание количества громкоговорителей I , доступных для воспроизведения. Таким образом, оно может устанавливаться независимо от параметров визуальной камеры или устройства
45 отображения.

Например, для равноотстоящих громкоговорителей, действительное значение усиления диффузного звука $Q \in [0, 1/\sqrt{I}]$ в формуле (2а) выбирается в блоке 202 выбора усиления

на основе параметра масштабирования β . Цель использования усиления диффузного звука состоит в том, чтобы ослаблять диффузный звук в зависимости от коэффициента масштабирования, например, масштабирование увеличивает DRR воспроизводимого сигнала. Это достигается посредством понижения Q для более большого β . Фактически, увеличение масштаба означает, что угол раскрытия камеры становится более малым, например, естественным акустическим соответствием будет более направленный микрофон, который захватывает меньше диффузного звука.

Чтобы имитировать этот эффект, один вариант осуществления может, например, использовать функцию усиления, показанную на фиг. 8. Фиг. 8 иллюстрирует пример функции усиления диффузного звука $q(\beta)$.

В других вариантах осуществления, функция усиления определяется другим образом. Окончательный диффузный звук $Y_{diff,i}(k, n)$ для i -ого канала громкоговорителя достигается посредством декоррелирования $Y_{diff}(k, n)$, например, согласно формуле (2b).

В последующем, рассматривается акустическое масштабирование на основе направлений DOA и расстояний.

Согласно некоторым вариантам осуществления, сигнальный процессор 105 может, например, быть сконфигурирован с возможностью принимать информацию расстояния, при этом сигнальный процессор 105 может, например, быть сконфигурирован с возможностью генерировать каждый выходной аудиосигнал из упомянутых одного или более выходных аудиосигналов в зависимости от информации расстояния.

Некоторые варианты осуществления используют обработку для совместимого акустического масштабирования, которое основывается как на оцененном DOA $\phi(k, n)$, так и на значении расстояния $r(k, n)$. Концепции этих вариантов осуществления также могут применяться, чтобы выравнивать записанную акустическую сцену с видео без масштабирования, где источники не располагаются на одном и том же расстоянии, как ранее предполагалось, доступная информация расстояния $r(k, n)$ обеспечивает нам возможность создавать эффект акустического размытия для источников звука, которые не проявляются отчетливыми в визуальном изображении, например, для источников, которые не располагаются на фокальной плоскости камеры.

Чтобы обеспечивать совместимое воспроизведение звука, например, акустическое масштабирование, с размытием для источников, расположенных на разных расстояниях, усиления $G_A(k, n)$ и Q могут регулироваться в формуле (2a), как изображено в модуле 103 модификации сигналов из фиг. 2, на основе двух оцененных параметров, именно $\phi(k, n)$ и $r(k, n)$, и в зависимости от коэффициента масштабирования β . Если никакое масштабирование не вовлекается, β может устанавливаться на $\beta=1$.

Параметры $\phi(k, n)$ и $r(k, n)$ могут, например, оцениваться в модуле 102 оценки параметров, как описано выше. В этом варианте осуществления, усиление прямого звука $G_A(k, n)$ определяется (например, посредством выбора в блоке 201 выбора усиления) на основе DOA и информации расстояния из одной или более функций усиления прямого звука $g_{i,j}(k, n)$ (которые могут, например, вычисляться в модуле 104 вычисления функций усиления). Аналогично, как описано для вариантов осуществления выше, усиление диффузного звука Q может, например, выбираться в блоке 202 выбора усиления от функции усиления диффузного звука $q(\beta)$, например, вычисленной в модуле 104 вычисления функций усиления на основе коэффициента масштабирования β .

В других вариантах осуществления, усиление прямого звука $G_A(k, n)$ и усиление диффузного звука Q вычисляются модулем 103 модификации сигналов без вычисления

сначала соответствующих функций усиления и затем выбора усилений.

Чтобы описать воспроизведение акустической сцены и акустическое масштабирование для источников звука на разных расстояниях, ссылка делается на фиг. 9. Параметры, обозначенные на фиг. 9, являются аналогичными параметрам, описанным выше.

5 На фиг. 9, источник звука располагается в положении P на расстоянии $R(k, n)$ до оси x . Расстояние r , которое может, например, быть зависящим от (k, n) (зависящим от времени-частоты: $r(k, n)$), обозначает расстояние между положением источника и фокальной плоскостью (левой вертикальной линией, проходящей через g). Следует отметить, что некоторые системы с автофокусировкой являются способными

10 обеспечивать g , например, расстояние до фокальной плоскости.

DOA прямого звука из точки обзора массива микрофонов обозначается посредством $\phi'(k, n)$. В отличие от других вариантов осуществления, не предполагается, что все источники расположены на одном и том же расстоянии g от объектива камеры. Таким образом, например, положение P может иметь произвольное расстояние $R(k, n)$ до оси

15 x .

Если источник не располагается на фокальной плоскости, источник будет проявляться размытым на видео. Более того, варианты осуществления основываются на обнаружении, что, если источник располагается в любом положении на пунктирной линии 910, он будет появляться в том же положении $x_b(k, n)$ на видео. Однако варианты

20 осуществления основываются на обнаружении, что оцененное DOA $\phi'(k, n)$ прямого звука изменяется, если источник перемещается вдоль пунктирной линии 910. Другими словами, на основе обнаружений, используемых вариантами осуществления, если источник перемещается параллельно оси y , оцененное DOA $\phi'(k, n)$ будет изменяться, в то время как x_b (и, таким образом, DOA $\phi_b(k, n)$, из которого звук должен

25 воспроизводиться) остается одним и тем же. Следовательно, если оцененное DOA $\phi'(k, n)$ передается стороне дальнего конца и используется для воспроизведения звука, как описано в предыдущих вариантах осуществления, то акустическое и визуальное изображение более не выровнены, если источник изменяет свое расстояние $R(k, n)$.

30 Чтобы компенсировать этот эффект и достигать совместимого воспроизведения звука, оценка DOA, например, выполняемая в модуле 102 оценки параметров, оценивает DOA прямого звука, как если бы источник располагался на фокальной плоскости в положении P . Это положение представляет проекцию P на фокальную плоскость. Соответствующее DOA обозначается посредством $\phi(k, n)$ на фиг. 9 и используется на

35 стороне дальнего конца для совместимого воспроизведения звука, аналогично тому, как в предыдущих вариантах осуществления. (Модифицированное) DOA $\phi(k, n)$ может вычисляться из оцененного (исходного) DOA $\phi'(k, n)$ на основе геометрических рассмотрений, если r и g являются известными.

Например, на фиг. 9, сигнальный процессор 105 может, например, вычислять $\phi(k, n)$ из $\phi'(k, n)$, r и g согласно:

$$40 \quad \phi = \arctan\left(\frac{\tan \phi' \cdot (r + g)}{g}\right).$$

Таким образом, согласно одному варианту осуществления, сигнальный процессор 105 может, например, быть сконфигурирован с возможностью принимать исходный

45 азимутальный угол $\phi'(k, n)$ направления прибытия, которое является направлением прибытия компонент прямых сигналов упомянутых двух или более входных аудиосигналов, и сконфигурирован с возможностью дополнительно принимать информацию расстояния, и может, например, быть сконфигурирован с возможностью

дополнительно принимать информацию расстояния l . Сигнальный процессор 105 может, например, быть сконфигурирован с возможностью вычислять модифицированный азимутальный угол $\phi(k, n)$ направления прибытия в зависимости от азимутального угла исходного направления прибытия $\phi'(k, n)$ и в зависимости от информации расстояния l и g . Сигнальный процессор 105 может, например, быть сконфигурирован с возможностью генерировать каждый выходной аудиосигнал из упомянутых одного или более из выходных аудиосигналов в зависимости от азимутального угла модифицированного направления прибытия $\phi(k, n)$.

Требуемая информация расстояния может оцениваться как описано выше (расстояние g фокальной плоскости может получаться из системы объектива или информации автофокуса). Следует отметить, что, например, в этом варианте осуществления, расстояние $l(k, n)$ между источником и фокальной плоскостью передается стороне дальнего конца вместе с (отображенным) DOA $\phi(k, n)$.

Более того, по аналогии с визуальным масштабированием, источники, лежащие на большом расстоянии l от фокальной плоскости, не проявляются отчетливыми в изображении. Этот эффект является хорошо известным в оптике как так называемая глубина поля (DOF), которая определяет диапазон расстояний источников, которые проявляются приемлемо отчетливо в визуальном изображении.

Пример кривой DOF как функции расстояния l изображен на фиг. 10(a). Фиг. 10 иллюстрирует примерные фигуры для глубины поля (фиг. 10(a)), для обрезающей частоты фильтра нижних частот (фиг. 10(b)), и для временной задержки в мс для повторного прямого звука (фиг. 10(c)).

На фиг. 10(a), источники на малом расстоянии от фокальной плоскости являются все еще отчетливыми, тогда как источники на более больших расстояниях (либо ближе, либо дополнительно дальше от камеры) проявляются как размытые. Таким образом, согласно одному варианту осуществления, соответствующие источники звука размываются, чтобы их визуальное и акустическое изображения были совместимыми.

Чтобы выводить усиления $G_i(k, n)$ и Q в (2a), которые реализуют акустическое размытие и совместимое воспроизведение пространственного звука, учитывается угол, под которым источник, расположенный в $P(\phi, l)$, будет появляться на устройстве отображения. Размытый источник будет отображаться под

$$\tan \varphi_b(k, n) \equiv \beta c \tan \varphi(k, n), \quad (30)$$

где c является параметром калибровки, $\beta \geq 1$ является управляемым пользователем коэффициентом масштабирования, $\varphi(k, n)$ является (отображенным) DOA, например, оцененным в модуле 102 оценки параметров. Как упомянуто ранее, усиление прямого звука $G_i(k, n)$ в таких вариантах осуществления может, например, вычисляться из множества функций усиления прямого звука $g_{i,j}$. В частности, могут, например, использоваться две функции усиления $g_{i,1}(\varphi(k, n))$ и $g_{i,2}(l(k, n))$, при этом первая функция усиления зависит от DOA $\varphi(k, n)$, и при этом вторая функция усиления зависит от расстояния $l(k, n)$. Усиление прямого звука $G_i(k, n)$ может вычисляться как:

$$G_i(k, n) \equiv g_{i,1}(\varphi(k, n)) g_{i,2}(l(k, n)), \quad (31)$$

$$g_{i,1}(\varphi) \equiv p_{b,i}(\varphi) w_b(\varphi), \quad (32)$$

$$g_{i,2}(l) \equiv b(l), \quad (33)$$

где $p_{b,i}(\varphi)$ обозначает функцию усиления панорамирования (чтобы гарантировать,

что звук воспроизводится из правильного направления), где $w_b(\phi)$ является оконной функцией усиления (чтобы гарантировать, что прямой звук ослабляется, если источник не является видимым на видео), и где $b(r)$ является функцией размытия (чтобы размывать источники акустически, если они не располагаются на фокальной плоскости).

5 Следует отметить, что все функции усиления могут определяться в зависимости от частоты (что пропускается здесь для краткости). Следует дополнительно отметить, что в этом варианте осуществления усиление прямого звука G_r находится посредством выбора и умножения усилений от двух разных функций усиления, как показано в формуле (32).

10 Обе функции усиления $p_{b,k}(\phi)$ и $w_b(\phi)$ определяются аналогично тому, как описано выше. Например, они могут вычисляться, например, в модуле 104 вычисления функций усиления, например, с использованием формул (26) и (27), и они остаются фиксированными, пока коэффициент масштабирования β не изменяется. Подробное описание этих двух функций было обеспечено выше. Функция размытия $b(r)$ возвращает комплексные усиления, которые вызывают размытие, например, перцепционное рассеивание, источника, и, таким образом, вся функция усиления g_r также будет обычно возвращать комплексное число. Для простоты, в последующем, размытие обозначается как функция расстояния до фокальной плоскости $b(r)$.

20 Эффект размытия может получаться как выбранный один или комбинация из следующих эффектов размытия: низкочастотная фильтрация, добавление задержанного прямого звука, ослабление прямого звука, временное сглаживание и/или рассеивание DOA. Таким образом, согласно одному варианту осуществления, сигнальный процессор 105 может, например, быть сконфигурирован с возможностью генерировать упомянутые
25 один или более выходных аудиосигналов посредством выполнения низкочастотной фильтрации, или посредством добавления задержанного прямого звука, или посредством выполнения ослабления прямого звука, или посредством выполнения временного сглаживания, или посредством выполнения рассеивания направления прибытия.

Низкочастотная фильтрация: В системах технического зрения, неотчетливое
30 визуальное изображение может получаться посредством низкочастотной фильтрации, которая эффективно осуществляет слияние соседних пикселей в визуальном изображении. По аналогии, эффект акустического размытия может получаться посредством низкочастотной фильтрации прямого звука с обрезавшей частотой, выбранной на основе оцененного расстояния источника до фокальной плоскости r . В
35 этом случае, функция размытия $b(r, k)$ возвращает усиления фильтра нижних частот для частоты k и расстояния r . Иллюстративная кривая для обрезавшей частоты фильтра нижних частот первого порядка для частоты дискретизации, равной 16 кГц, показана на фиг. 10(b). Для малых расстояний r , обрезавшая частота является близкой к частоте Найквиста, и, таким образом, почти никакая низкочастотная фильтрация фактически
40 не выполняется. Для значений более большого расстояния, обрезавшая частота уменьшается до тех пор, когда она выравнивается на 3 кГц, где акустическое изображение является достаточно размытым.

Добавление задержанного прямого звука: Чтобы делать акустическое изображение источника менее резким, мы можем декоррелировать прямой звук, например,
45 посредством повторения ослабления прямого звука после некоторой задержки τ (например, между 1 и 30 мс). Такая обработка может, например, выполняться согласно комплексной функции усиления из формулы (34):

$$b(r, k) = 1 + \alpha(r) e^{-j\omega\tau(r)}, (34)$$

где α обозначает усиление ослабления для повторного звука и τ является задержкой, после которой прямой звук повторяется. Иллюстративная кривая задержки (в мс) показана на фиг. 10(с). Для малых расстояний, задержанный сигнал не повторяется и α устанавливается на нуль. Для более больших расстояний, временная задержка
 5 увеличивается с увеличением расстояния, что вызывает перцепционное рассеивание акустического источника.

Ослабление прямого звука: Источник также может восприниматься как размытый, когда прямой звук ослабляется посредством постоянного коэффициента. В этом случае $b(r)=\text{const}<1$. Как упомянуто выше, функция размытия $b(r)$ может состоять из любого
 10 из упомянутых эффектов размытия или как комбинация этих эффектов. В дополнение, может использоваться альтернативная обработка, которая размывает источник.

Временное сглаживание: Сглаживание прямого звука по времени может, например, использоваться, чтобы перцепционно размывать акустический источник. Это может достигаться посредством сглаживания огибающей извлеченного прямого сигнала по
 15 времени.

Рассеивание DOA: Другой способ, чтобы делать акустический источник менее резким, состоит в воспроизведении исходного сигнала из диапазона направлений вместо воспроизведения только из оцененного направления. Это может достигаться посредством рандомизации угла, например, посредством взятия случайного угла из
 20 гауссовского распределения, центрированного около оцененного ϕ . Увеличение дисперсии такого распределения, и, таким образом, расширение возможного диапазона DOA, увеличивает восприятие размытия.

Аналогично тому, как описано выше, вычисление функции усиления диффузного звука $q(\beta)$ в модуле 104 вычисления функций усиления, может, в некоторых вариантах
 25 осуществления, требовать только знание количества громкоговорителей I , доступных для воспроизведения. Таким образом, функция усиления диффузного звука $q(\beta)$ может, в таких вариантах осуществления, устанавливаться, как требуется для применения. Например, для равноотстоящих громкоговорителей, действительное значение усиления
 30 диффузного звука $Q \in [0, 1/\sqrt{I}]$ в формуле (2а) выбирается в блоке 202 выбора усиления на основе параметра масштабирования β . Цель использования усиления диффузного звука состоит в том, чтобы ослаблять диффузный звук в зависимости от коэффициента масштабирования, например, масштабирование увеличивает DRR воспроизводимого сигнала. Это достигается посредством понижения Q для более большого β . Фактически, увеличение масштаба означает, что угол раскрытия камеры становится более малым,
 35 например, естественным акустическим соответствием будет более направленный микрофон, который захватывает меньше диффузного звука. Чтобы имитировать этот эффект, мы можем использовать, например, функцию усиления, показанную на фиг. 8. Ясно, что, функция усиления также может определяться различным образом.
 40 Необязательно, окончательный диффузный звук $Y_{diff,i}(k, n)$ для i -ого канала громкоговорителя получается посредством декоррелирования $Y_{diff}(k, n)$, полученного в формуле (2b).

Теперь, рассматриваются варианты осуществления, которые реализуют применение к слуховым аппаратам и вспомогательным слуховым устройствам. Фиг. 11 иллюстрирует
 45 такое применение к слуховым аппаратам.

Некоторые варианты осуществления относятся к бинауральным слуховым аппаратам. В этом случае, предполагается, что каждый слуховой аппарат оснащен, по меньшей мере, одним микрофоном и что информация может обмениваться между упомянутыми

двумя слуховыми аппаратами. Вследствие некоторой потери слуха, человек с нарушением слуха может испытывать трудности сосредоточения (например, концентрирования на звуках, идущих из конкретной точки или направления) на требуемом звуке или звуках. Чтобы помогать мозгу человека с нарушением слуха обрабатывать звуки, которые воспроизводятся слуховыми аппаратами, акустическое изображение делается совместимым с точкой фокуса или направлением пользователя слуховых аппаратов. Предполагается, что точка фокуса или направление является предварительно определенной, определяемой пользователем, или определяемой интерфейсом мозг-машина. Такие варианты осуществления обеспечивают, что требуемые звуки (которые предполагается, что прибывают из точки фокуса или направления фокуса) и нежелательные звуки проявляются пространственно разделенными.

В таких вариантах осуществления, направления прямых звуков могут оцениваться разными способами. Согласно одному варианту осуществления, направления определяются на основе интерауральных различий уровней (ILD) и/или интерауральных временных различий (ITD), которые определяются с использованием обоих слуховых аппаратов (см. [15] и [16]).

Согласно другим вариантам осуществления, направления прямых звуков на левой стороне и правой стороне оцениваются независимо с использованием слухового аппарата, который оснащен, по меньшей мере, двумя микрофонами (см. [17]). Оцененные направления могут использоваться на основе уровней звукового давления в левом и правом слуховом аппарате, или пространственной когерентности в левом и правом слуховом аппарате. Из-за эффекта затенения головой, разные средства оценки могут использоваться для разных частотных диапазонов (например, различия ILD при высоких частотах и различия ITD при низких частотах).

В некоторых вариантах осуществления, сигналы прямого и диффузного звуков могут, например, оцениваться с использованием вышеупомянутых способов информированной пространственной фильтрации. В этом случае, прямые и диффузные звуки, как принимаются в левом и правом слуховом аппарате, могут оцениваться отдельно (например, посредством изменения опорного микрофона), или левый и правый выходные сигналы могут генерироваться с использованием функции усиления для вывода левого и правого слухового аппарата, соответственно, аналогичным образом разные сигналы громкоговорителей или наушников получают в предыдущих вариантах осуществления.

Чтобы пространственно разделять требуемый и нежелательный звуки, может применяться акустическое масштабирование, описанное в вышеупомянутых вариантах осуществления. В этом случае, точка фокуса или направление фокуса определяет коэффициент масштабирования.

Таким образом, согласно одному варианту осуществления, может обеспечиваться слуховой аппарат или вспомогательное слуховое устройство, при этом слуховой аппарат или вспомогательное слуховое устройство содержит систему, как описано выше, при этом сигнальный процессор 105 вышеописанной системы определяет усиление прямого звука для каждого из упомянутых одного или более выходных аудиосигналов, например, в зависимости от направления фокуса или точки фокуса.

В одном варианте осуществления, сигнальный процессор 105 вышеописанной системы может, например, быть сконфигурирован с возможностью принимать информацию масштабирования. Сигнальный процессор 105 вышеописанной системы может, например, быть сконфигурирован с возможностью генерировать каждый выходной аудиосигнал из упомянутых одного или более выходных аудиосигналов в зависимости от оконной функции усиления, при этом оконная функция усиления зависит от

информации масштабирования. Используются такие же концепции, как описано со ссылкой на фиг. 7(a), 7(b) и 7(c).

Если аргумент оконной функции, в зависимости от направления фокуса или от точки фокуса, больше, чем нижний порог и меньше, чем верхний порог, оконная функция
5 усиления сконфигурирована с возможностью возвращать усиление окна, которое больше, чем любое усиление окна, возвращаемое оконной функцией усиления, если аргумент оконной функции меньше, чем нижний порог, или больше, чем верхний порог.

Например, в случае направления фокуса, направление фокуса может само быть аргументом оконной функции (и, таким образом, аргумент оконной функции зависит
10 от направления фокуса). В случае положения фокуса, аргумент оконной функции, может, например, выводиться из положения фокуса.

Аналогично, изобретение может применяться к другим носимым устройствам, которые включают в себя вспомогательные слуховые устройства или устройства, такие как Google Glass®. Следует отметить, что некоторые носимые устройства также
15 оснащаются одной или более камерами или датчиком ToF, который может использоваться, чтобы оценивать расстояние объектов до человека, носящего устройство.

Хотя некоторые аспекты были описаны в контексте устройства, должно быть ясно, что эти аспекты также представляют описание соответствующего способа, где блок
20 или устройство соответствует этапу способа или признаку этапа способа. Аналогично, аспекты, описанные в контексте этапа способа, также представляют описание соответствующего блока или элемента или признака соответствующего устройства.

Новый разложенный сигнал может сохраняться в цифровом запоминающем носителе или может передаваться посредством носителя передачи, такого как беспроводной
25 носитель передачи или проводной носитель передачи, такой как сеть Интернет.

В зависимости от некоторых требований вариантов осуществления, варианты осуществления изобретения могут осуществляться в аппаратном обеспечении или в программном обеспечении. Вариант осуществления может выполняться с
30 использованием цифрового запоминающего носителя, например, гибкого диска, DVD, CD, ROM, PROM, EPROM, EEPROM или флэш-памяти, имеющего электронно-читаемые сигналы управления, сохраненные на нем, которые взаимодействуют (или являются способными взаимодействовать) с программируемой компьютерной системой, чтобы выполнялся соответствующий способ.

Некоторые варианты осуществления согласно изобретению содержат нетранзиторный
35 носитель данных, имеющий электронно-читаемые сигналы управления, которые являются способными взаимодействовать с программируемой компьютерной системой, чтобы выполнялся один из способов, здесь описанных.

В общем, варианты осуществления настоящего изобретения могут осуществляться как компьютерный программный продукт с программным кодом, при этом
40 программный код является работоспособным для выполнения одного из способов, когда компьютерный программный продукт исполняется на компьютере. Программный код может, например, быть сохранен на машиночитаемом носителе.

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

Другими словами, один вариант осуществления нового способа является, поэтому, компьютерной программой, имеющей программный код для выполнения одного из
45 способов, здесь описанных, когда компьютерная программа исполняется на компьютере.

Один дополнительный вариант осуществления новых способов является, поэтому,

носителем данных (или цифровым запоминающим носителем, или компьютерно-читаемым носителем), содержащим, записанную на нем, компьютерную программу для выполнения одного из способов, здесь описанных.

5 Один дополнительный вариант осуществления нового способа является, поэтому, потоком данных или последовательностью сигналов, представляющим компьютерную программу для выполнения одного из способов, здесь описанных. Поток данных или последовательность сигналов может, например, быть сконфигурирован с возможностью передаваться посредством соединения передачи данных, например, посредством сети Интернет.

10 Один дополнительный вариант осуществления содержит средство обработки, например, компьютер, или программируемое логическое устройство, сконфигурированное с возможностью или выполненное с возможностью выполнять один из способов, здесь описанных.

15 Один дополнительный вариант осуществления содержит компьютер, имеющий, установленную на нем компьютерную программу для выполнения одного из способов, здесь описанных.

В некоторых вариантах осуществления, программируемое логическое устройство (например, программируемая пользователем вентильная матрица) может использоваться, чтобы выполнять некоторые или все из функциональных возможностей 20 способов, здесь описанных. В некоторых вариантах осуществления, программируемая пользователем вентильная матрица может взаимодействовать с микропроцессором, чтобы выполнять один из способов, здесь описанных. В общем, способы предпочтительно выполняются посредством любого аппаратного устройства.

25 Вышеописанные варианты осуществления являются всего лишь иллюстративными для принципов настоящего изобретения. Следует понимать, что модификации и изменения компоновок и подробностей, здесь описанных, должны быть ясны другим специалистам в данной области техники. Поэтому предполагается ограничение только посредством объема приложенной патентной формулы изобретения и не посредством конкретных подробностей, представленных посредством описания и объяснения 30 вариантов осуществления отсюда.

ССЫЛОЧНЫЕ ИСТОЧНИКИ

[1] Y. Ishigaki, M. Yamamoto, K. Totsuka, and N. Miyaji, "Zoom microphone", in Audio Engineering Society Convention 67, Paper 1713, October 1980.

35 [2] M. Matsumoto, H. Naono, H. Saitoh, K. Fujimura, and Y. Yasuno, "Stereo zoom microphone for consumer video cameras", Consumer Electronics, IEEE Transactions on, vol. 35, no. 4, pp. 759-766, November 1989. August 13, 2014

[3] T. van Waterschoot, W. J. Tirry, and M. Moonen, "Acoustic zooming by multi microphone sound scene manipulation", J. Audio Eng. Soc, vol. 61, no. 7/8, pp. 489-507, 2013.

40 [4] V. Pulkki, "Spatial sound reproduction with directional audio coding", J. Audio Eng. Soc, vol. 55, no. 6, pp. 503-516, June 2007.

[5] R. Schultz-Amling, F. Kuech, O. Thiergart, and M. Kallinger, "Acoustical zooming based on a parametric sound field representation", in Audio Engineering Society Convention 128, Paper 8120, London UK, May 2010.

45 [6] O. Thiergart, G. Del Galdo, M. Taseska, and E. Habets, "Geometry-based spatial sound acquisition using distributed microphone arrays", Audio, Speech, and Language Processing, IEEE Transactions on, vol. 21, no. 12, pp. 2583-2594, December 2013.

[7] K. Kowalczyk, O. Thiergart, A. Craciun, and E. A. P. Habets, "Sound acquisition in noisy and reverberant environments using virtual microphones", in Applications of Signal Processing

to Audio and Acoustics (WASPAA), 2013 IEEE Workshop on, October 2013.

[8] O. Thiergart and E. A. P. Habets, "An informed LCMV filter based on multiple instantaneous direction-of-arrival estimates", in Acoustics Speech and Signal Processing (ICASSP), 2013 IEEE International Conference on, 2013, pp. 659-663.

5 [9] O. Thiergart and E. A. P. Habets, "Extracting reverberant sound using a linearly constrained minimum variance spatial filter", Signal Processing Letters, IEEE, vol. 21, no. 5, pp. 630-634, May 2014.

[10] R. Roy and T. Kailath, "ESPRIT-estimation of signal parameters via rotational invariance techniques", Acoustics, Speech and Signal Processing, IEEE Transactions on, vol. 37, no. 7, pp. 984-995, July 1989.

10 [11] B. Rao and K. Hari, "Performance analysis of root-music", in Signals, Systems and Computers, 1988. Twenty-Second Asilomar Conference on, vol. 2, 1988, pp. 578-582.

[12] H. Teutsch and G. Elko, "An adaptive close-talking microphone array", in Applications of Signal Processing to Audio and Acoustics, 2001 IEEE Workshop on the, 2001, pp. 163-166.

15 [13] O. Thiergart, G. D. Galdo, and E. A. P. Habets, "On the spatial coherence in mixed sound fields and its application to signal-to-diffuse ratio estimation", The Journal of the Acoustical Society of America, vol. 132, no. 4, pp. 2337-2346, 2012.

[14] V. Pulkki, "Virtual sound source positioning using vector base amplitude panning", J. Audio Eng. Soc, vol. 45, no. 6, pp. 456-466, 1997.

20 [15] J. Blauert, Spatial hearing, 3rd ed. Hirzel-Verlag, 2001.

[16] T. May, S. van de Par, and A. Kohlrausch, "A probabilistic model for robust localization based on a binaural auditory front-end", IEEE Trans. Audio, Speech, Lang. Process., vol. 19, no. 1, pp. 1-13, 2011.

25 [17] J. Ahonen, V. Sivonen, and V. Pulkki, "Parametric spatial sound processing applied to bilateral hearing aids", in AES 45th International Conference, Mar. 2012.

(57) Формула изобретения

1. Устройство для генерирования одного или более выходных аудиосигналов, содержащее:

30 сигнальный процессор (105) и
 интерфейс (106) вывода,

 при этом сигнальный процессор (105) сконфигурирован с возможностью принимать сигнал прямых компонент, содержащий компоненты прямых сигналов из двух или более исходных аудиосигналов, при этом сигнальный процессор (105) сконфигурирован с возможностью принимать сигнал диффузных компонент, содержащий компоненты диффузных сигналов из упомянутых двух или более исходных аудиосигналов, и при этом сигнальный процессор (105) сконфигурирован с возможностью принимать информацию направления, при этом упомянутая информация направления зависит от направления прибытия компонент прямых сигналов из упомянутых двух или более исходных аудиосигналов,

40 при этом сигнальный процессор (105) сконфигурирован с возможностью генерировать один или более обработанных диффузных сигналов в зависимости от сигнала диффузных компонент,

 при этом для каждого выходного аудиосигнала из упомянутых одного или более выходных аудиосигналов сигнальный процессор (105) сконфигурирован с возможностью определять, в зависимости от направления прибытия, усиление прямого звука, которое является значением усиления, сигнальный процессор (105) сконфигурирован с возможностью применять упомянутое усиление прямого звука к сигналу прямых

компонент, чтобы получать обработанный прямой сигнал, и сигнальный процессор (105) сконфигурирован с возможностью комбинировать упомянутый обработанный прямой сигнал и один из упомянутых одного или более обработанных диффузных сигналов, чтобы генерировать упомянутый выходной аудиосигнал, и

5 при этом интерфейс (106) вывода сконфигурирован с возможностью выводить упомянутые один или более выходных аудиосигналов,

при этом сигнальный процессор (105) содержит модуль (104) вычисления функций усиления для вычисления одной или более функций усиления, при этом каждая функция усиления из упомянутых одной или более функций усиления содержит множество значений аргумента функции усиления, при этом возвращаемое значение функции усиления назначено каждому из упомянутых значений аргумента функции усиления, при этом, когда упомянутая функция усиления принимает одно из упомянутых значений аргумента функции усиления, упомянутая функция усиления сконфигурирована с возможностью возвращать возвращаемое значение функции усиления, которое
10 назначено упомянутому одному из упомянутых значений аргумента функции усиления,
и

при этом сигнальный процессор (105) дополнительно содержит модуль (103) модификации сигналов для выбора, в зависимости от направления прибытия, зависящего от направления значения аргумента из значений аргумента функции усиления для
20 функции усиления из упомянутых одной или более функций усиления, для получения возвращаемого значения функции усиления, которое назначено упомянутому зависящему от направления значению аргумента, от упомянутой функции усиления, и для определения значения усиления по меньшей мере одного из упомянутых одного или более выходных аудиосигналов в зависимости от упомянутого возвращаемого значения
25 функции усиления, полученного от упомянутой функции усиления.

2. Система для генерирования одного или более выходных аудиосигналов, содержащая:

устройство по п. 1 и

модуль (101) генерирования сигнала компонент,

30 при этом модуль (101) генерирования сигнала компонент сконфигурирован с возможностью принимать два или более входных аудиосигналов, которые являются упомянутыми двумя или более исходными аудиосигналами,

при этом модуль (101) генерирования сигнала компонент сконфигурирован с возможностью генерировать сигнал прямых компонент, содержащий компоненты
35 прямых сигналов из упомянутых двух или более исходных аудиосигналов, и

при этом модуль (101) генерирования сигнала компонент сконфигурирован с возможностью генерировать сигнал диффузных компонент, содержащий компоненты диффузных сигналов из упомянутых двух или более исходных аудиосигналов.

3. Система по п. 2,

40 в которой модуль (104) вычисления функций усиления сконфигурирован с возможностью генерировать таблицу поиска для каждой функции усиления из упомянутых одной или более функций усиления, при этом таблица поиска содержит множество записей, при этом каждая из записей таблицы поиска содержит одно из значений аргумента функции усиления и возвращаемое значение функции усиления,
45 которое назначено упомянутому значению аргумента функции усиления,

при этом модуль (104) вычисления функций усиления сконфигурирован с возможностью сохранять таблицу поиска каждой функции усиления в постоянной или непостоянной памяти и

при этом модуль (103) модификации сигналов сконфигурирован с возможностью получать возвращаемое значение функции усиления, которое назначено упомянутому зависящему от направления значению аргумента, посредством считывания упомянутого возвращаемого значения функции усиления из одной из упомянутых одной или более таблиц поиска, которые сохранены в памяти.

4. Система по п. 2,

в которой сигнальный процессор (105) сконфигурирован с возможностью определять два или более выходных аудиосигналов,

при этом модуль (104) вычисления функций усиления сконфигурирован с возможностью вычислять две или более функций усиления,

при этом для каждого выходного аудиосигнала из упомянутых двух или более выходных аудиосигналов модуль (104) вычисления функций усиления сконфигурирован с возможностью вычислять функцию усиления панорамирования, которая назначена упомянутому выходному аудиосигналу в качестве одной из упомянутых двух или более функций усиления, при этом модуль (103) модификации сигналов сконфигурирован с возможностью генерировать упомянутый выходной аудиосигнал в зависимости от упомянутой функции усиления панорамирования.

5. Система по п. 4,

в которой функция усиления панорамирования каждого из упомянутых двух или более выходных аудиосигналов имеет один или более глобальных максимумов, являющихся одним из значений аргумента функции усиления упомянутой функции усиления панорамирования, при этом для каждого из упомянутых одного или более глобальных максимумов упомянутой функции усиления панорамирования не существует никакое другое значение аргумента функции усиления, для которого упомянутая функция усиления панорамирования возвращает более большое возвращаемое значение функции усиления, чем для упомянутых глобальных максимумов, и

при этом для каждой пары из первого выходного аудиосигнала и второго выходного аудиосигнала из упомянутых двух или более выходных аудиосигналов по меньшей мере один из упомянутых одного или более глобальных максимумов функции усиления панорамирования первого выходного аудиосигнала отличается от любого из упомянутых одного или более глобальных максимумов функции усиления панорамирования второго выходного аудиосигнала.

6. Система по п. 4,

в которой для каждого выходного аудиосигнала из упомянутых двух или более выходных аудиосигналов модуль (104) вычисления функций усиления сконфигурирован с возможностью вычислять оконную функцию усиления, которая назначена упомянутому выходному аудиосигналу в качестве одной из упомянутых двух или более функций усиления,

при этом модуль (103) модификации сигналов сконфигурирован с возможностью генерировать упомянутый выходной аудиосигнал в зависимости от упомянутой оконной функции усиления и

при этом если значение аргумента упомянутой оконной функции усиления больше, чем нижний порог окна, и меньше, чем верхний порог окна, оконная функция усиления сконфигурирована с возможностью возвращать возвращаемое значение функции усиления, которое больше, чем любое возвращаемое значение функции усиления, возвращаемое упомянутой оконной функцией усиления, если значение аргумента оконной функции меньше, чем нижний порог, или больше, чем верхний порог.

7. Система по п. 6,

в которой оконная функция усиления каждого из упомянутых двух или более выходных аудиосигналов имеет один или более глобальных максимумов, являющихся одним из значений аргумента функции усиления упомянутой оконной функции усиления, при этом для каждого из упомянутых одного или более глобальных максимумов упомянутой оконной функции усиления не существует никакого другого значение аргумента функции усиления, для которого упомянутая оконная функция усиления возвращает более большое возвращаемое значение функции усиления, чем для упомянутых глобальных максимумов, и

при этом для каждой пары из первого выходного аудиосигнала и второго выходного аудиосигнала из упомянутых двух или более выходных аудиосигналов по меньшей мере один из упомянутых одного или более глобальных максимумов оконной функции усиления первого выходного аудиосигнала является равным одному из упомянутых одного или более глобальных максимумов оконной функции усиления второго выходного аудиосигнала.

8. Система по п. 6,

в которой модуль (104) вычисления функций усиления сконфигурирован с возможностью дополнительно принимать информацию ориентации, указывающую угловой сдвиг направления просмотра по отношению к направлению прибытия, и при этом модуль (104) вычисления функций усиления сконфигурирован с

возможностью генерировать функцию усиления панорамирования каждого из выходных аудиосигналов в зависимости от информации ориентации.

9. Система по п. 8, в которой модуль (104) вычисления функций усиления сконфигурирован с возможностью генерировать оконную функцию усиления каждого из выходных аудиосигналов в зависимости от информации ориентации.

10. Система по п. 6,

в которой модуль (104) вычисления функций усиления сконфигурирован с возможностью дополнительно принимать информацию масштабирования, при этом информация масштабирования указывает угол раскрытия камеры, и

при этом модуль (104) вычисления функций усиления сконфигурирован с

возможностью генерировать функцию усиления панорамирования каждого из выходных аудиосигналов в зависимости от информации масштабирования.

11. Система по п. 10, в которой модуль (104) вычисления функций усиления сконфигурирован с возможностью генерировать оконную функцию усиления каждого из выходных аудиосигналов в зависимости от информации масштабирования.

12. Система по п. 6,

в которой модуль (104) вычисления функций усиления сконфигурирован с возможностью дополнительно принимать параметр калибровки для выравнивания визуального изображения и акустического изображения, и

при этом модуль (104) вычисления функций усиления сконфигурирован с

возможностью генерировать функцию усиления панорамирования каждого из выходных аудиосигналов в зависимости от параметра калибровки.

13. Система по п. 12, в которой модуль (104) вычисления функций усиления сконфигурирован с возможностью генерировать оконную функцию усиления каждого из выходных аудиосигналов в зависимости от параметра калибровки.

14. Система по п. 2,

в которой модуль (104) вычисления функций усиления сконфигурирован с возможностью принимать информацию о визуальном изображении, и

при этом модуль (104) вычисления функций усиления сконфигурирован с

возможностью генерировать, в зависимости от информации о визуальном изображении, функцию размытия, возвращающую комплексные усиления, чтобы реализовать перцепционное рассеивание источника звука.

15. Способ для генерирования одного или более выходных аудиосигналов, содержащий:

прием сигнала прямых компонент, содержащего компоненты прямых сигналов из двух или более исходных аудиосигналов,

прием сигнала диффузных компонент, содержащего компоненты диффузных сигналов из упомянутых двух или более исходных аудиосигналов,

10 прием информации направления, при этом упомянутая информация направления зависит от направления прибытия компонент прямых сигналов из упомянутых двух или более исходных аудиосигналов,

генерирование одного или более обработанных диффузных сигналов в зависимости от сигнала диффузных компонент,

15 для каждого выходного аудиосигнала из упомянутых одного или более выходных аудиосигналов определение, в зависимости от направления прибытия, усиления прямого звука, применение упомянутого усиления прямого звука к сигналу прямых компонент, чтобы получать обработанный прямой сигнал, и комбинирование упомянутого обработанного прямого сигнала и одного из упомянутых одного или более
20 обработанных диффузных сигналов, чтобы генерировать упомянутый выходной аудиосигнал, и

вывод упомянутых одного или более выходных аудиосигналов,

при этом генерирование упомянутых одного или более выходных аудиосигналов содержит вычисление одной или более функций усиления, при этом каждая функция

25 усиления из упомянутых одной или более функций усиления содержит множество значений аргумента функции усиления, при этом возвращаемое значение функции усиления назначено каждому из упомянутых значений аргумента функции усиления, при этом, когда упомянутая функция усиления принимает одно из упомянутых значений аргумента функции усиления, упомянутая функция усиления сконфигурирована с
30 возможностью возвращать возвращаемое значение функции усиления, которое назначено упомянутому одному из упомянутых значений аргумента функции усиления, и

при этом генерирование упомянутых одного или более выходных аудиосигналов содержит выбор, в зависимости от направления прибытия, зависящего от направления
35 значения аргумента из значений аргумента функции усиления для функции усиления из упомянутых одной или более функций усиления, для получения возвращаемого значения функции усиления, которое назначено упомянутому зависящему от направления значению аргумента, от упомянутой функции усиления и для определения значения усиления по меньшей мере одного из упомянутых одного или более выходных
40 аудиосигналов в зависимости от упомянутого возвращаемого значения функции усиления, полученного от упомянутой функции усиления.

16. Способ по п. 15, в котором способ дополнительно содержит:

прием двух или более входных аудиосигналов, которые являются упомянутыми двумя или более исходными аудиосигналами,

45 генерирование сигнала прямых компонент, содержащего компоненты прямых сигналов из упомянутых двух или более исходных аудиосигналов, и

генерирование сигнала диффузных компонент, содержащего компоненты диффузных сигналов из упомянутых двух или более исходных аудиосигналов.

17. Машиночитаемый носитель, имеющий компьютерную программу для осуществления способа по п. 15 или 16, когда исполняется на компьютере или сигнальном процессоре.

5

10

15

20

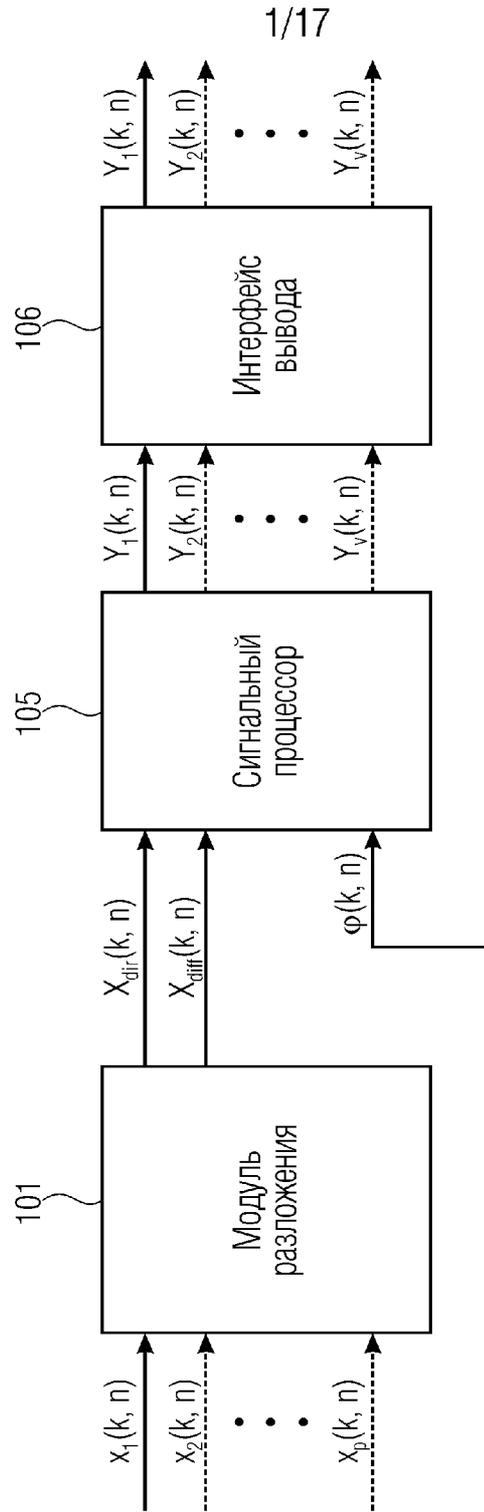
25

30

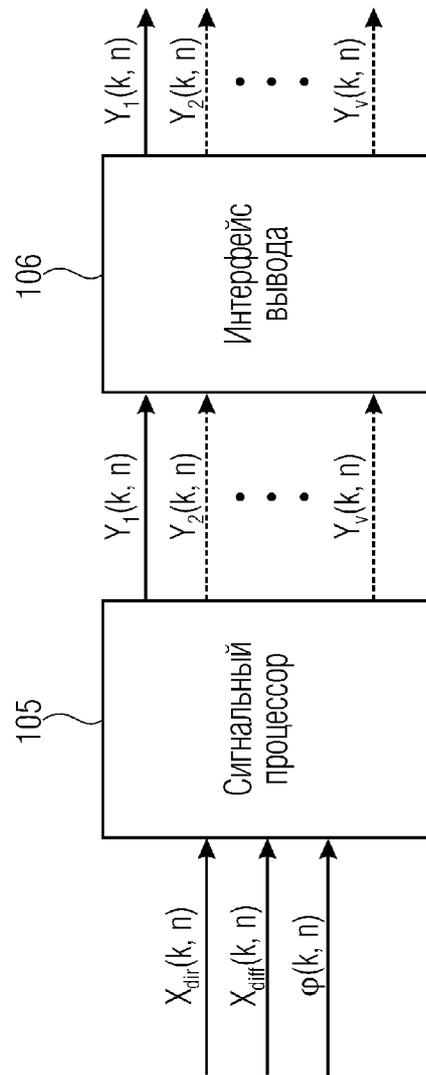
35

40

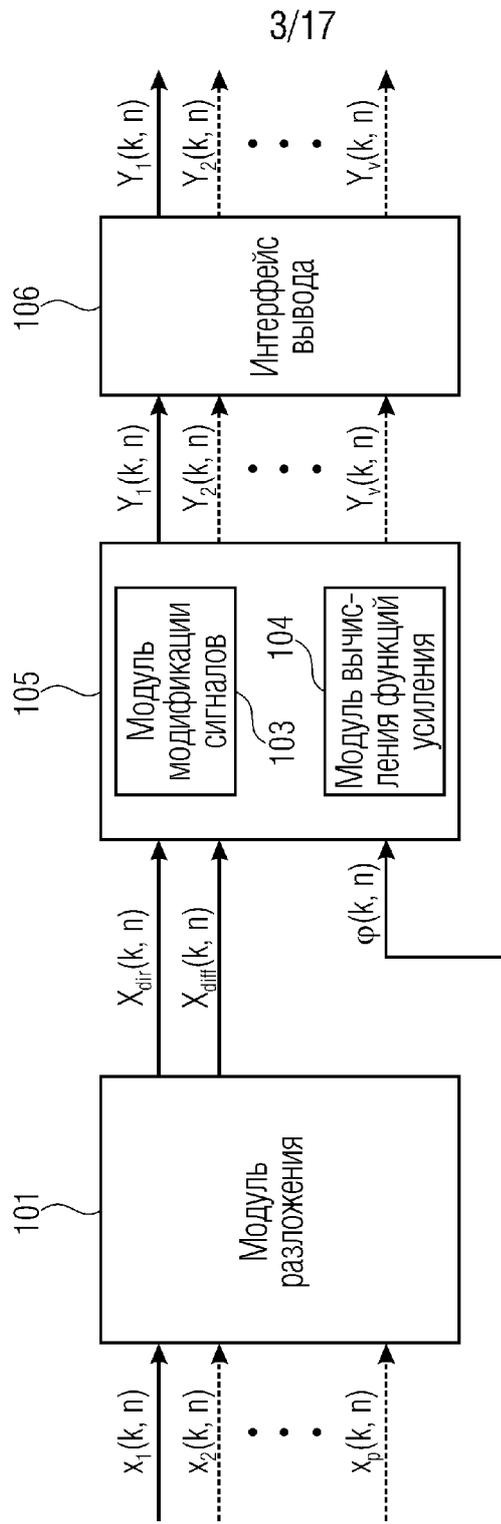
45



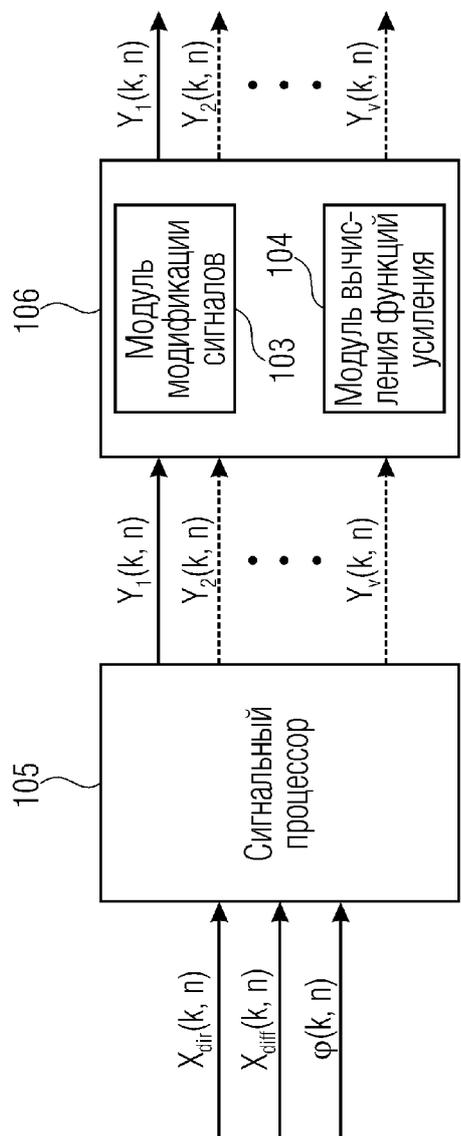
ФИГ. 1А



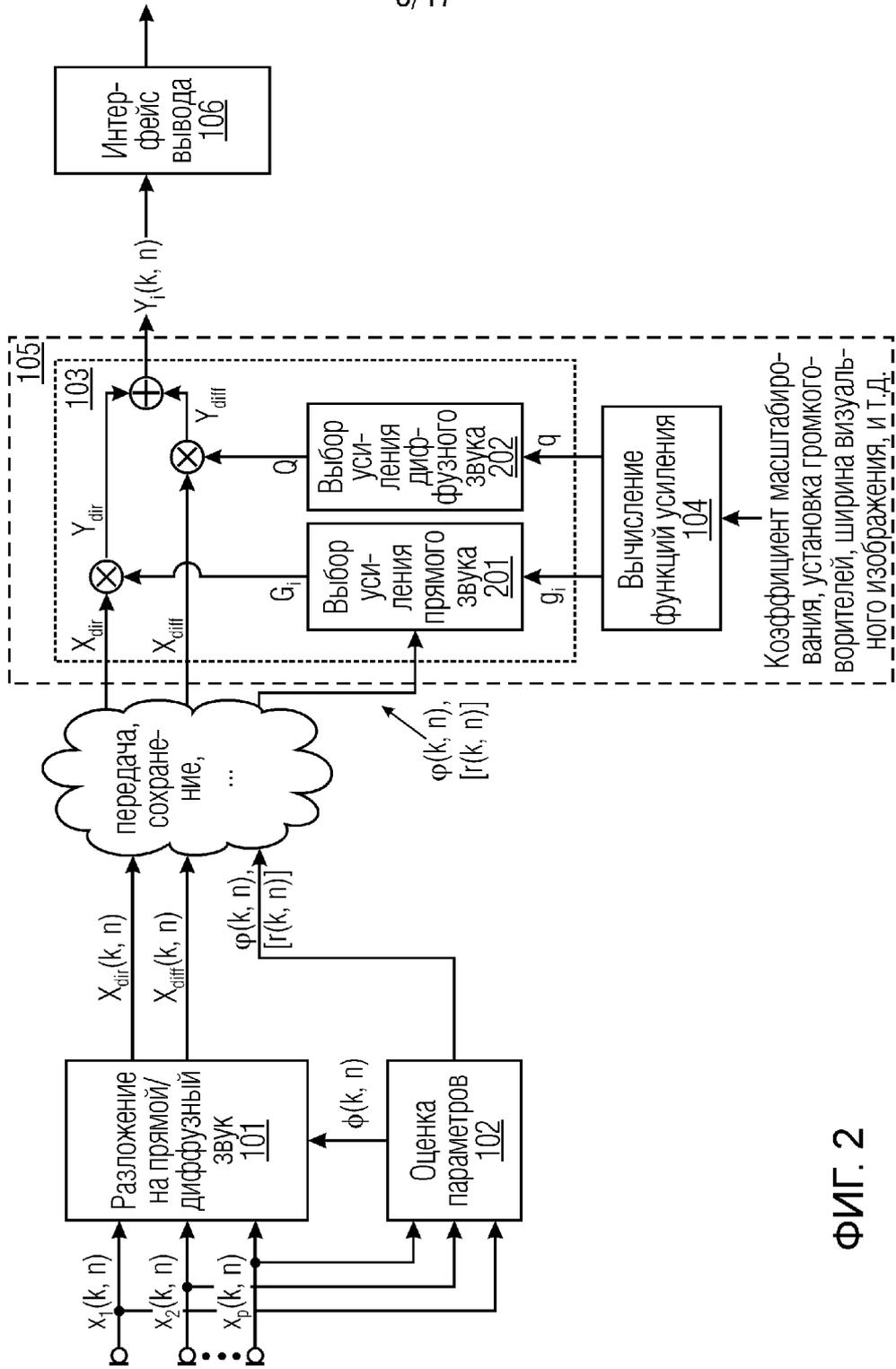
ФИГ. 1В



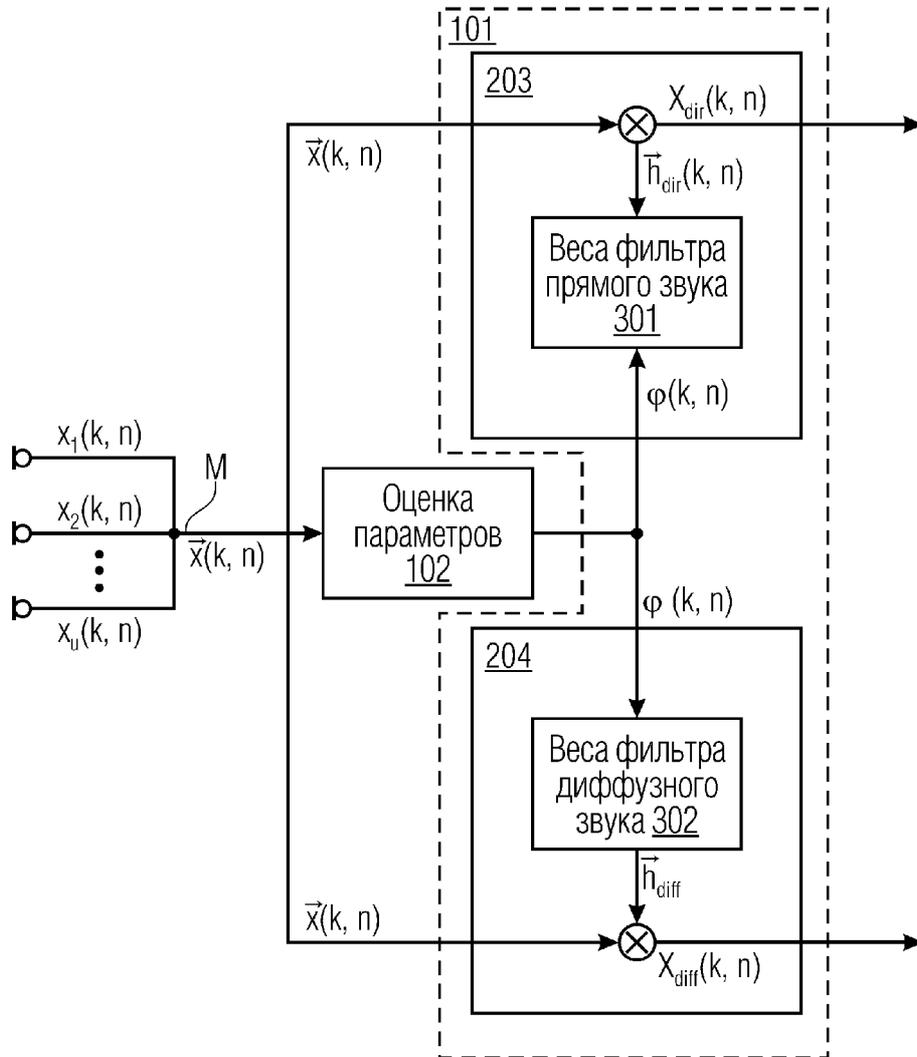
ФИГ. 1С



ФИГ. 1D

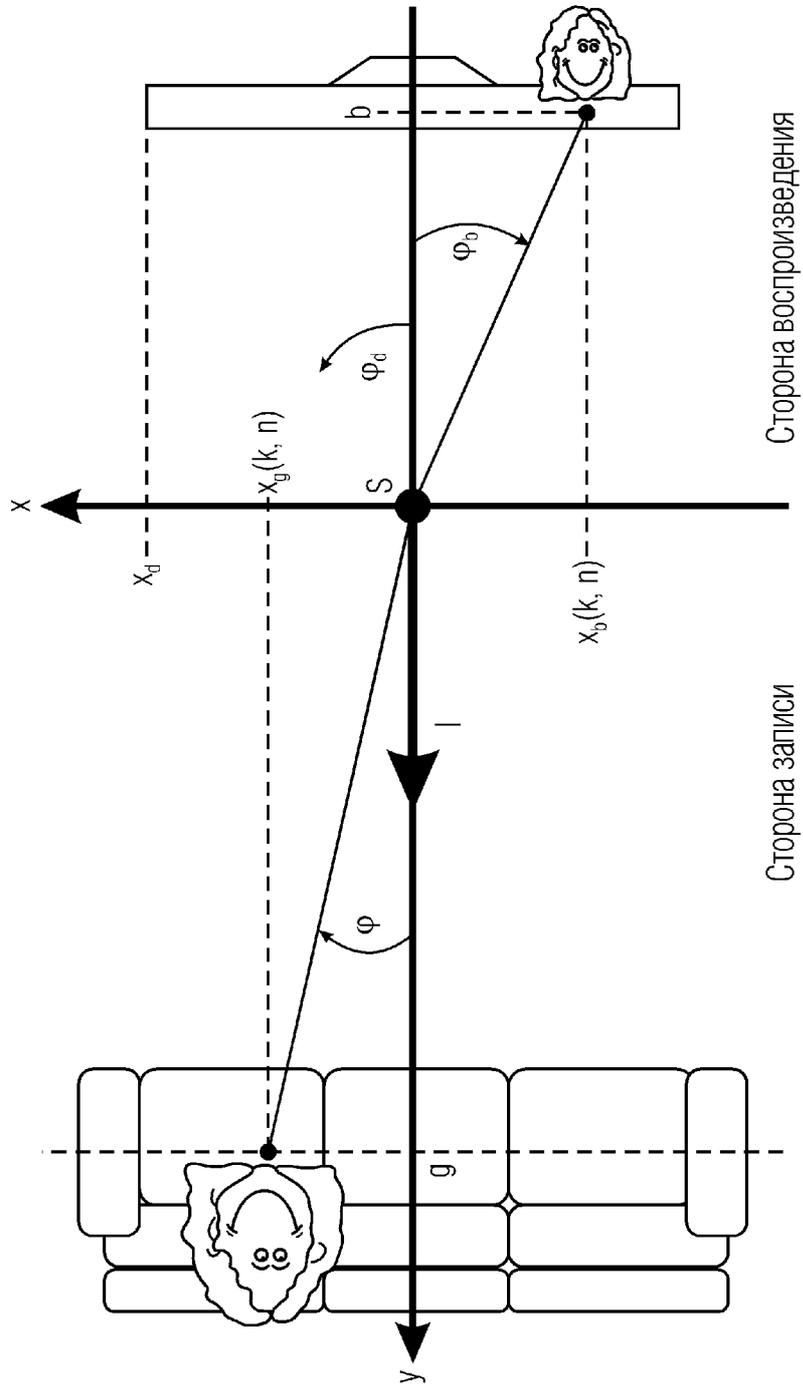


ФИГ. 2



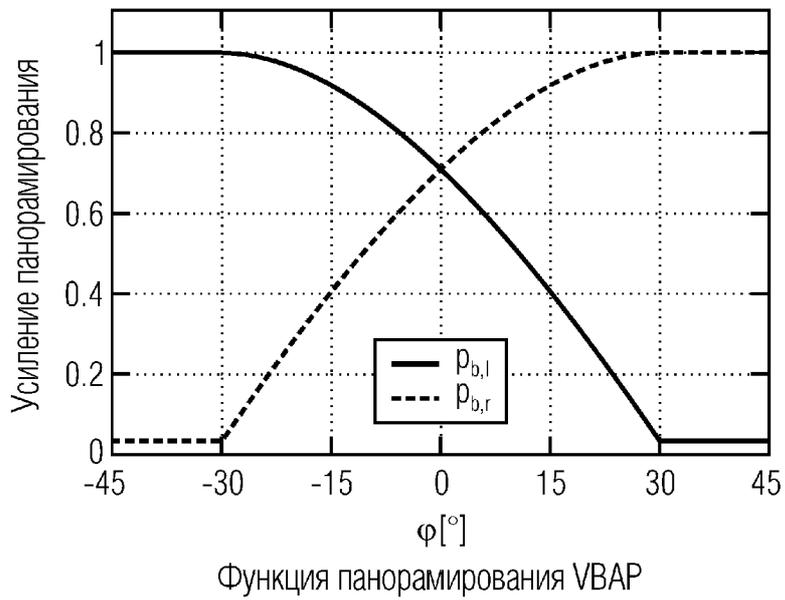
ФИГ. 3

7/17

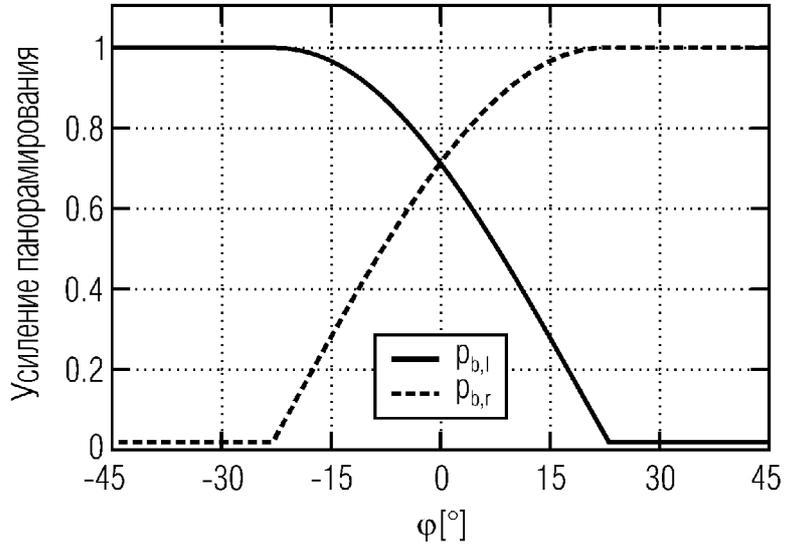


ФИГ. 4

8/17

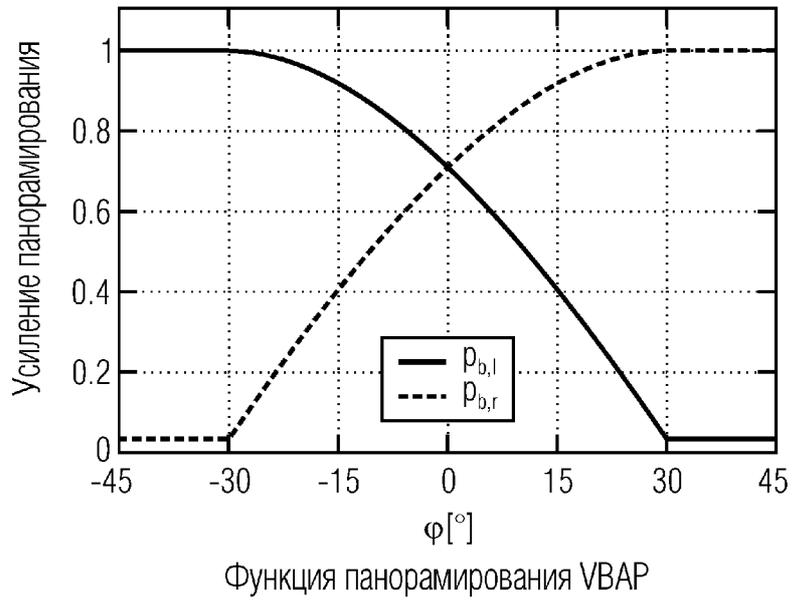


ФИГ. 5А

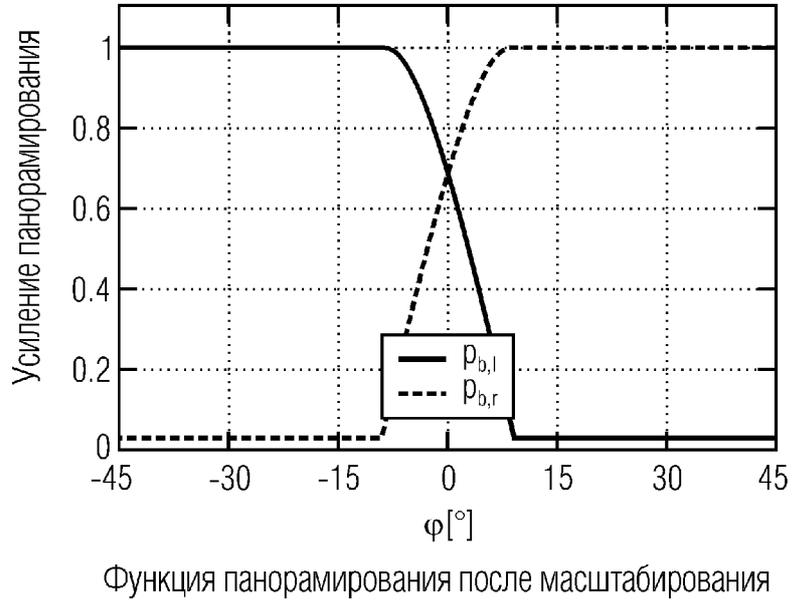


ФИГ. 5В

9/17

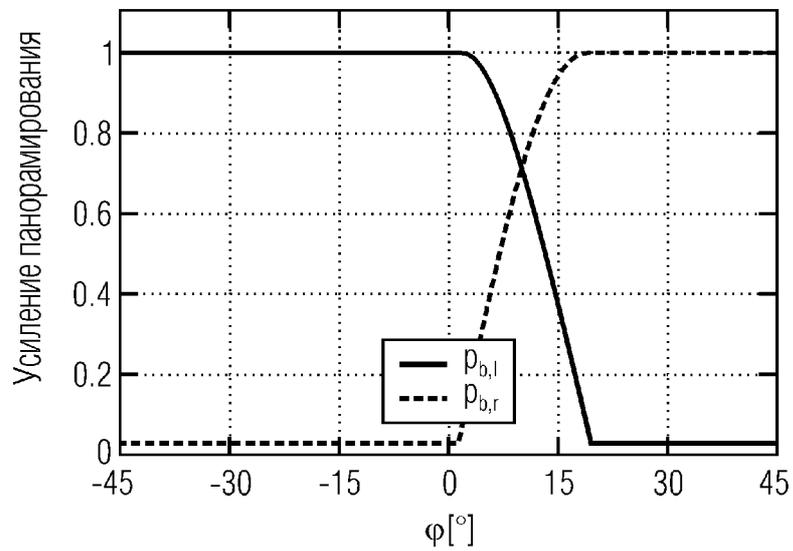


ФИГ. 6А



ФИГ. 6В

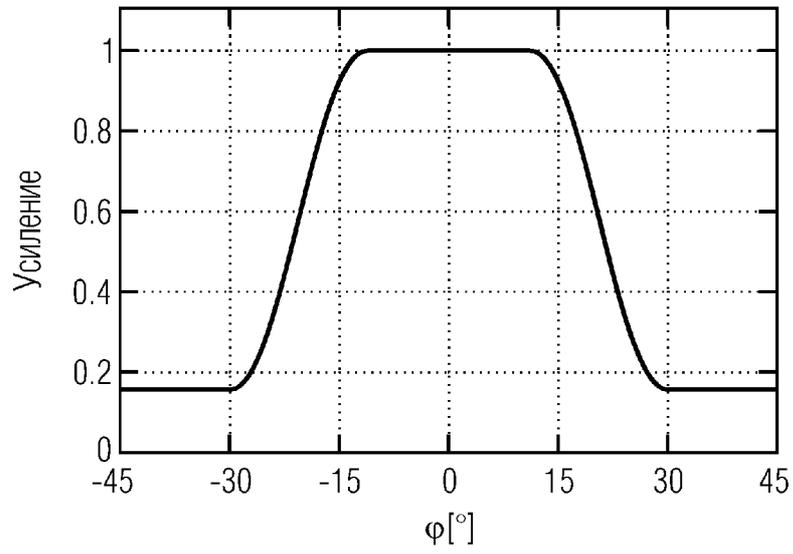
10/17



Функция панорамирования после масштабирования со сдвигом

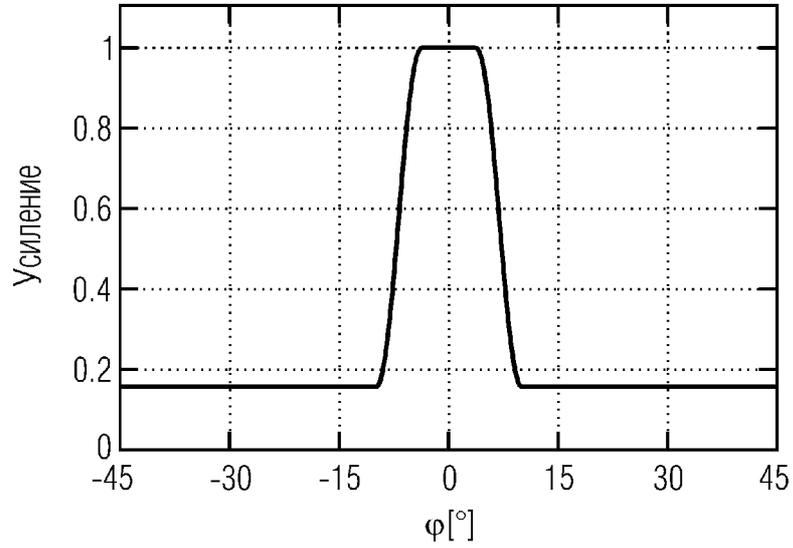
ФИГ. 6С

11/17



Оконная функция

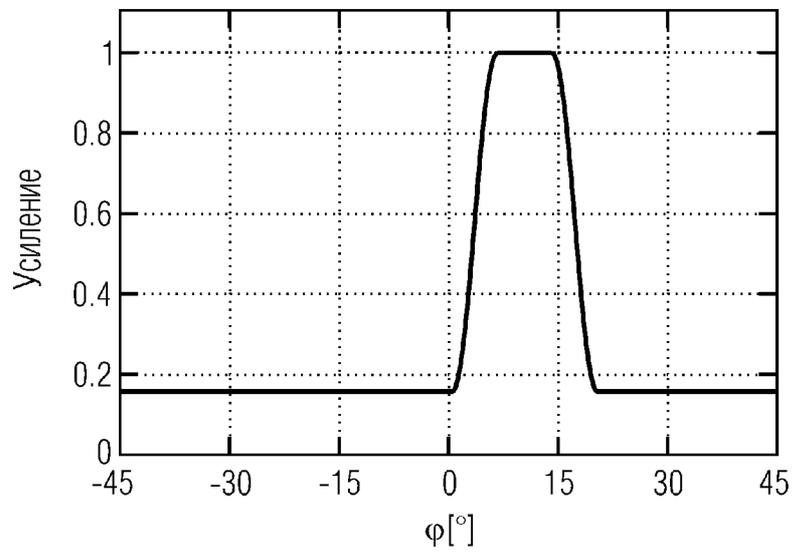
ФИГ. 7А



Оконная функция после масштабирования

ФИГ. 7В

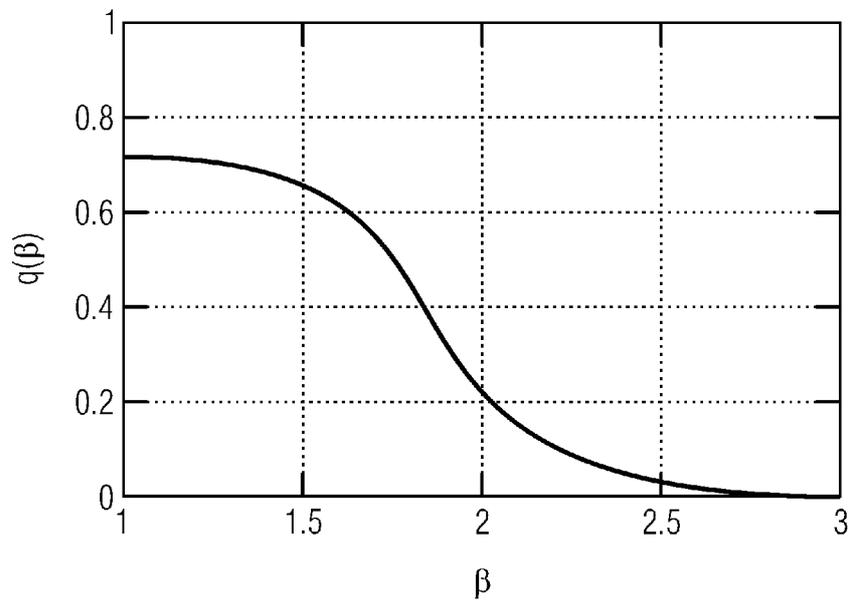
12/17



Оконная функция после масштабирования со сдвигом

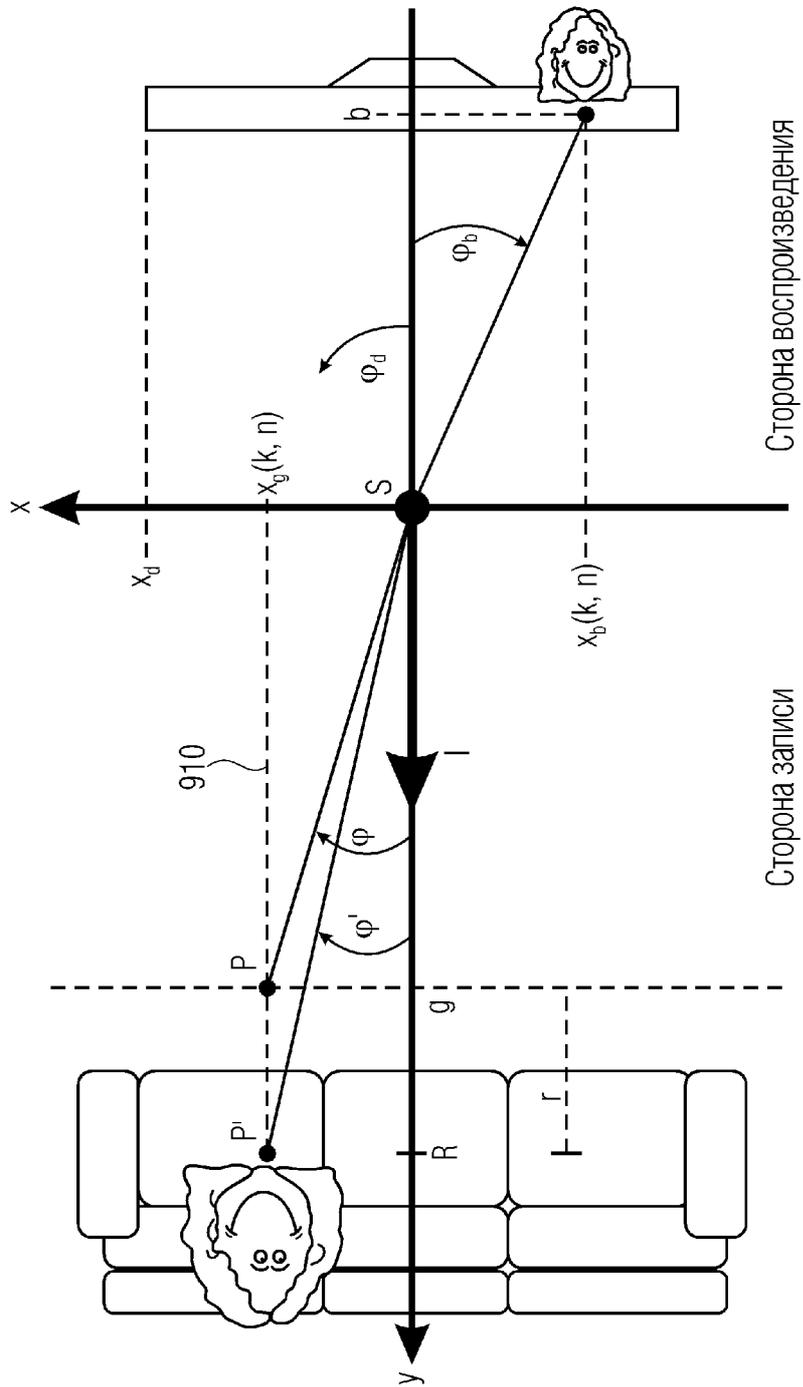
ФИГ. 7С

13/17



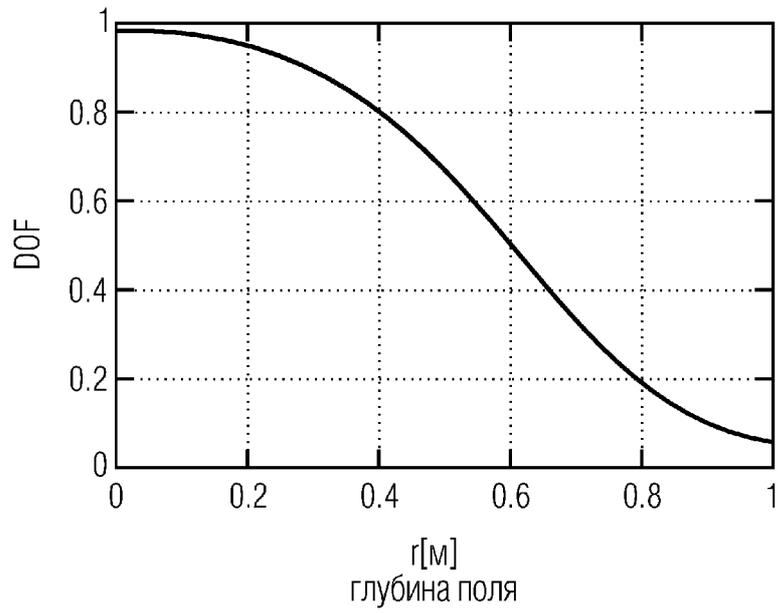
ФИГ. 8

14/17

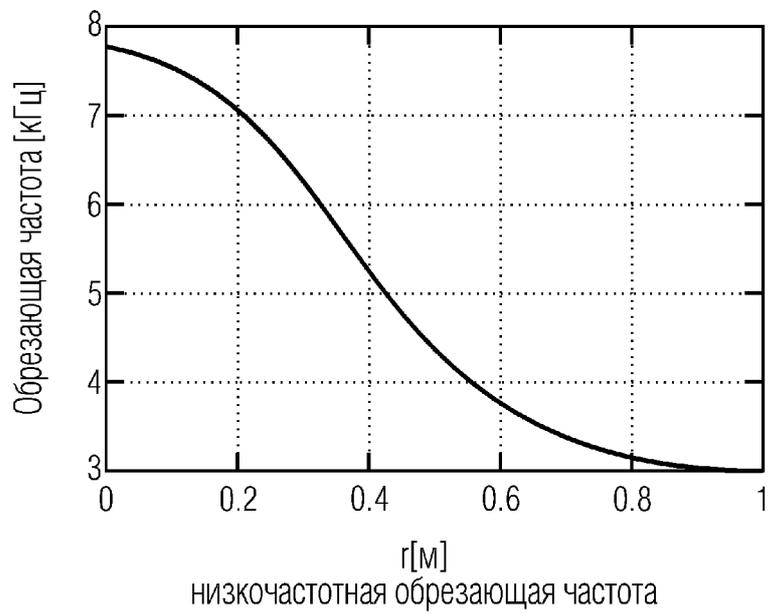


ФИГ. 9

15/17

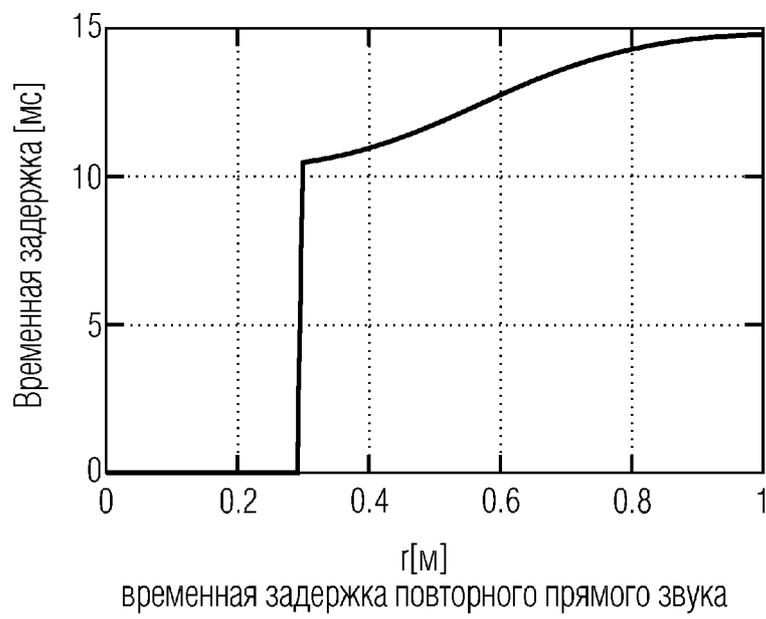


ФИГ. 10А



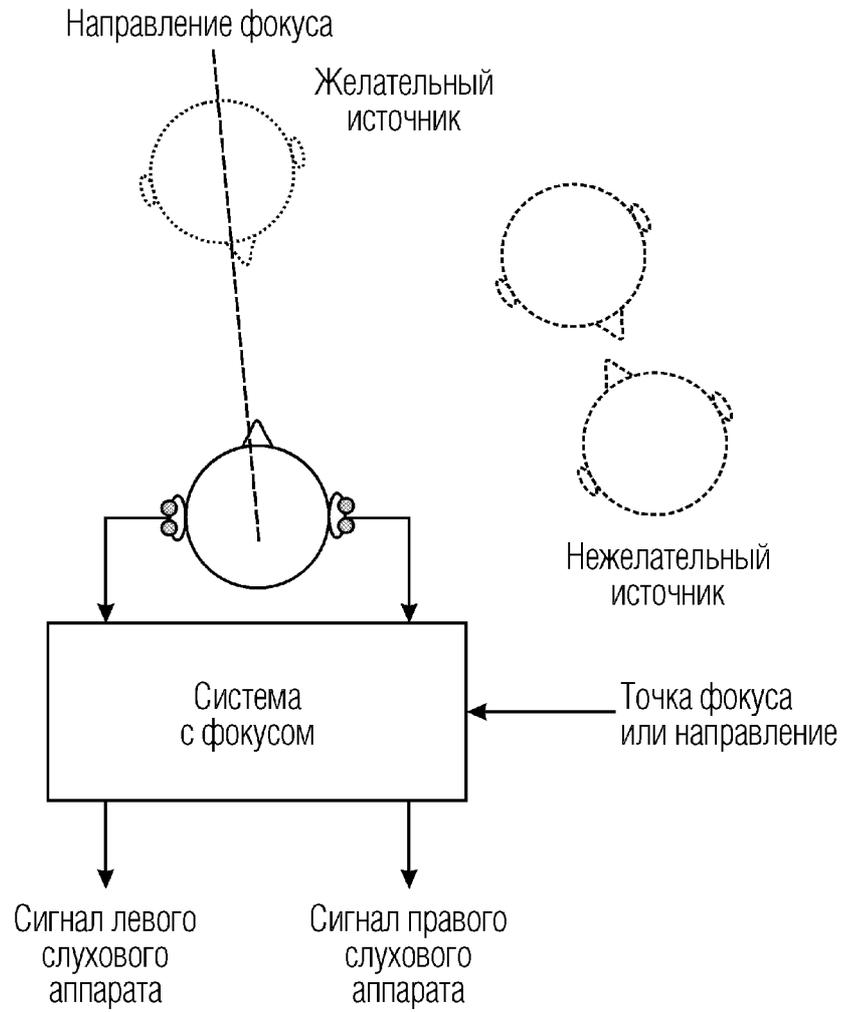
ФИГ. 10В

16/17



ФИГ. 10С

17/17



ФИГ. 11