



ФЕДЕРАЛЬНАЯ СЛУЖБА  
ПО ИНТЕЛЛЕКТУАЛЬНОЙ СОБСТВЕННОСТИ

## (12) ОПИСАНИЕ ИЗОБРЕТЕНИЯ К ПАТЕНТУ

(52) СПК

G10L 21/02 (2006.01); G10L 19/008 (2006.01)

(21)(22) Заявка: 2016127191, 02.12.2014

(24) Дата начала отсчета срока действия патента:  
02.12.2014

Дата регистрации:  
11.01.2018

Приоритет(ы):

(30) Конвенционный приоритет:  
11.12.2013 EP 13196672.3;  
20.02.2014 EP 14156014.4

(45) Опубликовано: 11.01.2018 Бюл. № 2

(85) Дата начала рассмотрения заявки РСТ на  
национальной фазе: 11.07.2016

(86) Заявка РСТ:  
EP 2014/076252 (02.12.2014)

(87) Публикация заявки РСТ:  
WO 2015/086377 (18.06.2015)

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

(72) Автор(ы):

ТИРГАРТ Оливер (DE),  
ХАБЕТС Эмануэль (NL)

(73) Патентообладатель(и):

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

(56) Список документов, цитированных в отчете  
о поиске: US 2009/0299742 A1, 03.12.2009. US  
2012/0128160 A1, 24.05.2012. US 2009/0252355  
A1, 08.10.2009. RU 2343562 C1, 10.01.2009. US  
2008/0069366 A1, 20.03.2008. WO 2008/034221  
A1, 27.03.2008. US 2012/0063608 A1, 15.03.2012.

## (54) ИЗВЛЕЧЕНИЕ РЕВЕРБЕРИРУЮЩЕГО ЗВУКА С ИСПОЛЬЗОВАНИЕМ МИКРОФОННЫХ МАССИВОВ

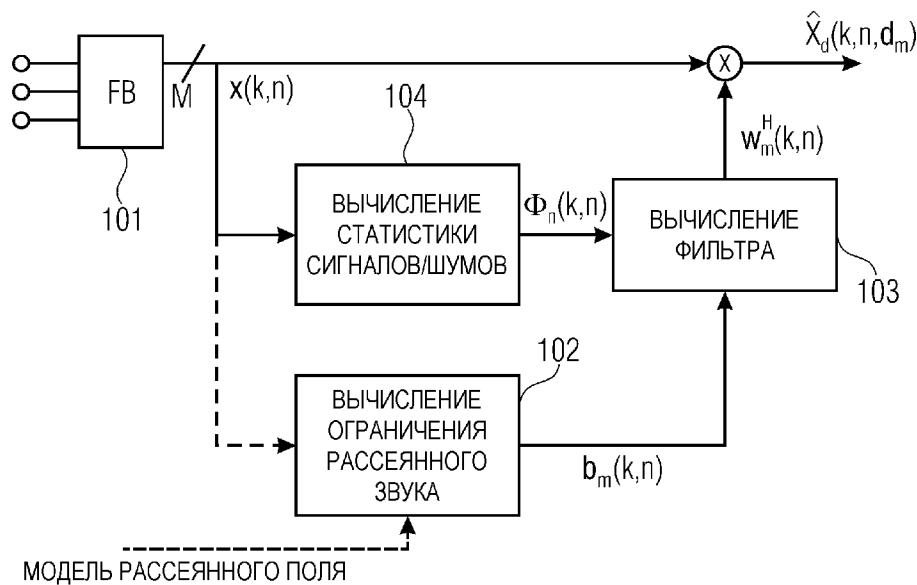
(57) Реферат:

Изобретение относится к области акустического анализа. Технический результат – обеспечение захвата рассеянного звука, поступающего на микрофонный массив со всех направлений, путем получения лучшей диаграммы направленности фильтра рассеянного звука. Способ содержит этап, на котором оценивают пространственную когерентность между первым участком рассеянного звука в первом микрофонном сигнале и вторым участком рассеянного звука во втором микрофонном сигнале. Первый микрофонный сигнал

захватывается первым микрофоном, а второй микрофонный сигнал захватывается вторым микрофоном, который расположен на расстоянии от первого микрофона. Способ дополнительно содержит этап, на котором определяют линейное ограничение для коэффициентов фильтрации фильтра рассеянного звука, причем линейное ограничение основано на пространственной когерентности. Способ также содержит этап, на котором вычисляют по меньшей мере одно из статистики сигналов и статистики шумов в отношении первого микрофонного сигнала и

второго микрофонного сигнала. Способ также содержит этап, на котором определяют коэффициенты фильтрации фильтра рассеянного звука путем решения оптимизационной задачи в

отношении по меньшей мере одного из статистики сигналов и статистики шумов, с учетом линейного ограничения для коэффициентов фильтрации. 3 н. и 12 з.п. ф-лы, 8 ил.



ФИГ. 3



FEDERAL SERVICE  
FOR INTELLECTUAL PROPERTY

**(12) ABSTRACT OF INVENTION**

(52) CPC

*G10L 21/02 (2006.01); G10L 19/008 (2006.01)*(21)(22) Application: **2016127191, 02.12.2014**(24) Effective date for property rights:  
**02.12.2014**Registration date:  
**11.01.2018**

Priority:

(30) Convention priority:  
**11.12.2013 EP 13196672.3;**  
**20.02.2014 EP 14156014.4**(45) Date of publication: **11.01.2018 Bull. № 2**(85) Commencement of national phase: **11.07.2016**(86) PCT application:  
**EP 2014/076252 (02.12.2014)**(87) PCT publication:  
**WO 2015/086377 (18.06.2015)**

Mail address:

**129090, Moskva, ul. B.Spasskaya, 25, stroenie 3,**  
**OOO "Yuridicheskaya firma Gorodisskij i Partnery"**

(72) Inventor(s):

**TIRGART Oliver (DE),**  
**KHABETS Emanuel (NL)**

(73) Proprietor(s):

**FRAUNKHOFER-GEZELLSCHAFT TSUR**  
**FERDERUNG DER ANGEVANDTEN**  
**FORSHUNG E.F. (DE)****(54) EXTRACTION OF REVERBERATIVE SOUND USING MICROPHONE MASSIVES**

(57) Abstract:

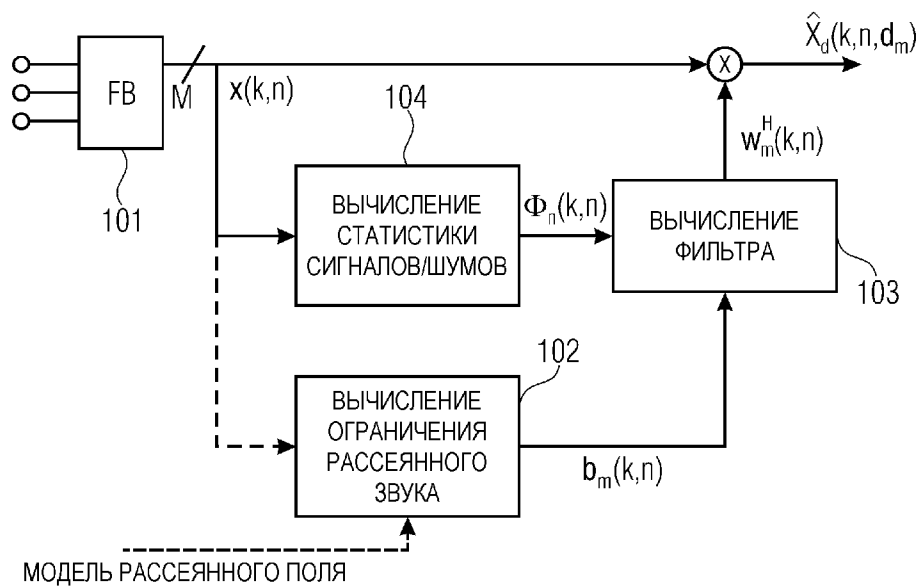
FIELD: physics.

SUBSTANCE: method contains the stages, on which the spatial coherence between the first phase of the scattered sound in the first microphone signal and the second phase of the ambient sound in the second microphone signal is estimated. The first microphone signal is captured by the first microphone, and the second microphone signal is captured by the second microphone, which is located at a distance from the first microphone. The method additionally contains a stage, at which the linear restrictions of the filtration coefficients of the ambient sound filter is determined, and the linear constraint is based on spatial coherence. The method also comprises the step of calculating at

least one of the signal statistics and noise statistics with respect to the first microphone signal and the second microphone signal. The method also contains a stage, at which the coefficients of the filter filtration of ambient sound filter is determined by solving the optimisation problem against at least one of signal statistics and noise statistics, given the linear restrictions of the filtration coefficients.

EFFECT: ensuring the capture of scattered sound coming to the microphone array from all directions, by obtaining a better directional pattern of the scattered sound filter.

15 cl, 8 dwg



ФИГ. 3

Нижеследующее изобретение относится к области акустического анализа, записи пространственного звука, обработке сигналов микрофонных массивов, и пространственной фильтрации. Некоторые варианты осуществления настоящего изобретения относятся к способу, который может использоваться для определения

5 коэффициентов фильтрации фильтра рассеянного звука, т.е. фильтра для извлечения рассеянного звука (реверберирующего звука) из записей с использованием микрофонного массива. Некоторые варианты осуществления относятся к соответствующей компьютерной программе. Некоторые варианты осуществления относятся к устройству, которое может быть использовано для определения

10 коэффициентов фильтрации фильтра рассеянного звука.

Захват звука с использованием микрофонного массива в реверберирующих средах обычно направлен на захват прямого звука от источников звука, при ослаблении шума и реверберации. Для многих применений было бы выгодным, если бы можно было также извлекать реверберирующий звук, при подавлении прямого звука и шума.

15 Например, в пространственном воспроизведении звука [Pulkki2007, Thiergart2013, Kowalczyk2013], реверберация, присутствующая на стороне записи, должна быть воспроизведена на стороне воспроизведения для воссоздания необходимого стереофонического эффекта. Кроме того, при обеспечении оценки реверберирующего звука, мы можем вычислить параметры, такие как отношение сигнал-реверберация или

20 мощность реверберирующего звука, которые представляют важную информацию для различных других применений.

В то время как оценка компонентов прямого звука (например, с использованием разделения источников, дереверберации, или подавления шума) хорошо освещена в литературе, существует лишь небольшое количество методик для извлечения

25 реверберирующего звука. Обычно реверберация моделируется в виде (изменяющегося во времени) рассеянного звукового поля. Для извлечения рассеянного звука, в последнее время использовались одноканальные фильтры (например, в [Pulkki2007, Thiergart2013]), которые имеют плохую производительность, когда активно множество источников, или для сигналов, подобных сигналам переходных процессов. Лучшая

30 производительность может быть достигнута с использованием многоканальных фильтров (например, [Kowalczyk2013, Thiergart2013b]). К сожалению, существующие в настоящее время многоканальные фильтры являются неоптимальными и не обеспечивают подходящую диаграмму направленности для захвата рассеянного звука.

Поэтому желательно обеспечить фильтр рассеянного звука, имеющий улучшенную

35 производительность в отношении извлечения рассеянного звука и/или подавления прямого звука. Также, может быть желательным, чтобы фильтр рассеянного звука имел диаграмму направленности, которая является в высшей степени всенаправленной, за исключением направлений поступления компонентов прямого звука. В высшей степени всенаправленная характеристика направленности является желательной,

40 поскольку рассеянный звук поступает на микрофонный массив со всех направлений.

Предложен способ, который содержит определение линейного ограничения для коэффициентов фильтрации фильтра рассеянного звука. Линейное ограничение основано на пространственной когерентности между первым участком рассеянного звука в первом микрофонном сигнале и вторым участком рассеянного звука во втором

45 микрофонном сигнале. Первый микрофонный сигнал захвачен первым микрофоном, а второй микрофонный сигнал захвачен вторым микрофоном, расположенным на расстоянии от первого микрофона, известным способом. Способ также содержит вычисление по меньшей мере одного из направления поступления по меньшей мере

одного прямого звука, статистики сигналов в отношении первого и второго микрофонных сигналов и статистики шумов в отношении первого и второго микрофонных сигналов. Способ дополнительно содержит определение коэффициентов фильтрации фильтра рассеянного звука путем решения оптимизационной задачи в отношении по меньшей мере одного из направления поступления по меньшей мере одного прямого звука, статистики сигналов и статистики шумов, с учетом линейного ограничения для коэффициентов фильтрации.

Варианты осуществления обеспечивают компьютерную программу для реализации вышеописанного способа, когда обеспечивается ее исполнение на компьютере или процессоре обработки сигналов.

Дополнительные варианты осуществления обеспечивают устройство, содержащее вычислитель линейного ограничения, выполненный с возможностью определения линейного ограничения для коэффициентов фильтрации фильтра рассеянного звука. Линейное ограничение основано на пространственной когерентности между первым участком рассеянного звука в первом микрофонном сигнале и вторым участком рассеянного звука во втором микрофонном сигнале. Первый микрофонный сигнал захвачен или был захвачен первым микрофоном, а второй микрофонный сигнал захвачен или был захвачен вторым микрофоном, расположенным на расстоянии от первого микрофона, известным способом. Устройство также содержит вычислитель статистики, выполненный с возможностью вычисления по меньшей мере одного из направления поступления по меньшей мере одного прямого звука, статистики сигналов в отношении первого и второго микрофонных сигналов и статистики шумов в отношении первого и второго микрофонных сигналов и второго микрофонного сигнала. Устройство дополнительно содержит вычислитель коэффициентов фильтрации, выполненный с возможностью определения коэффициентов фильтрации фильтра рассеянного звука, путем решения оптимизационной задачи в отношении по меньшей мере одного из направления поступления по меньшей мере одного прямого звука, статистики сигналов и статистики шумов, с учетом линейного ограничения для коэффициентов фильтрации.

Варианты осуществления основаны на понимании того, что фильтр рассеянного звука может быть определен с учетом по меньшей мере одного линейного ограничения, которое относится к участкам рассеянного звука микрофонных сигналов.

Далее варианты осуществления настоящего изобретения описаны более подробно со ссылкой на чертежи, на которых:

Фиг. 1 показывает схематичную блок-схему методики для извлечения рассеянного звука с использованием одноканального фильтра;

Фиг. 2 показывает схематичную блок-схему методики для извлечения рассеянного звука с использованием многоканального фильтра;

Фиг. 3 показывает схематичную блок-схему предлагаемого изобретения согласно первому примеру для реализации;

Фиг. 4 показывает схематичную блок-схему предлагаемого изобретения согласно второму примеру для реализации;

Фиг. 5 показывает схематичную блок-схему предлагаемого изобретения согласно третьему примеру для реализации;

Фиг. 6 показывает пример диаграммы направленности захвата, полученной в результате применения методики для извлечения рассеянного звука с использованием фильтра, согласно фиг. 2;

Фиг. 7 показывает пример диаграммы направленности захвата, полученной в результате применения методики для извлечения рассеянного звука с использованием

фильтра, согласно фиг. 4; и

Фиг. 8 схематично показывает микрофонный массив и разные звуковые сигналы, которые захвачены микрофонным массивом.

Термины «прямой звук» и «рассеянный звук» определены следующим образом.

5 Прямой звук: звуки, которые поступают в микрофоны, главным образом, по конкретному выраженному направлению. Прямой звук может представлять собой, например, звук, поступающий прямо от источника звука к микрофону, или отдельное отражение помещения. Прямые звуки могут быть, например, плоскими волнами или сферическими волнами с конкретным направлением поступления. Когда направление  
10 поступления прямого звука известно, можно вычислить относительную передаточную функцию прямого звука между микрофонами при условии, что геометрия микрофонов известна.

Рассеянный звук: звук, который поступает в микрофоны со всех направлений. Рассеянный звук может представлять собой, например, позднюю реверберацию,  
15 присутствующую в помещении. Обычно, никакие выраженные направления поступления не могут быть связаны с рассеянным звуком (изотропное звуковое поле), т.е. звук поступает с одинаковой средней мощностью со всех направлений. Кроме того, необходимо предположить, что относительные передаточные функции рассеянного звука между микрофонами являются случайными и ненаблюдаемыми. Однако, средние  
20 относительные передаточные функции рассеянного звука между микрофонами обычно известны для конкретных расположений микрофонов и моделей рассеянного поля или могут быть измерены.

Нижеследующие подсекции являются кратким изложением существующих методик для извлечения рассеянного звука (или реверберирующего звука) из микрофонной  
25 записи. Далее,  $M$  означает количество используемых микрофонов. Мы предполагаем, что все микрофонные сигналы были преобразованы в частотно-временную область, где  $k$  является индексом частоты, а  $n$  является индексом времени (следует отметить, что обычно фильтры могут также применяться во временной области). Микрофоны захватывают  $L$  плоских волн (называемых прямым звуком), распространяющихся в  
30 рассеянном поле. DOA  $l$ -ой плоской волны представлено единичным вектором нормали  $\mathbf{n}_l(k, n)$ . В частотно-временной области, сигнал  $m$ -го (всенаправленного) микрофона может быть представлен следующим образом:

$$35 \quad X_m(k, n) = \sum_{l=1}^L X_l(k, n, \mathbf{d}_m) + X_d(k, n, \mathbf{d}_m) + X_n(k, n, \mathbf{d}_m)$$

Здесь,  $X_l(k, n)$  является звуковым давлением  $l$ -ой плоской волны,  $X_d(k, n, \mathbf{d}_m)$  является  
рассеянным звуком,  $X_n(k, n, \mathbf{d}_m)$  является стационарным шумом (например, собственным шумом или фоновым шумом), и  $\mathbf{d}_m$  является вектором, описывающим положение  
40 микрофона ( $m$ -го микрофона) в данной системе координат.

Целью настоящего изобретения является оценка  $X_d(k, n, \mathbf{d}_m)$  в положении  $\mathbf{d}_m$ .

Одноканальные фильтры извлекают рассеянный звук из единственного микрофонного сигнала ( $M = 1$ ). Такие фильтры используются, например, в Направленном звуковом  
45 кодировании [Pulkki2007] или в Виртуальном микрофоне [Thiergart2013].

Оценка рассеянного звука определяется путем умножения одного из микрофонных сигналов, например, микрофонного сигнала первого микрофона  $X_1(k, n)$ , на фильтр  $H(k, n)$ , например,

$$\hat{X}_d(k, n, \mathbf{d}_m) = X_1(k, n)H(k, n)$$

Обычно, фильтр  $H(k, n)$  является фильтром Винера, который может быть представлен в виде

$$H(k, n) = \frac{\phi_d(k, n)}{\phi_d(k, n) + \phi_n(k, n)}$$

где  $\phi_d$  является мощностью рассеянного звука, а  $\phi_n$  является мощностью плоских волн и стационарного шума. В некоторых применениях, квадратный корень из фильтра Винера (т.е. квадратный корень из  $H$ ) используется вместо фильтра Винера. Следует отметить, что для вычисления  $H(k, n)$ , необходимо оценить мощность  $\phi_d$  и  $\phi_n$ . Для этой цели, мы можем рассмотреть, например, отношение сигнал-рассеяние (SDR), которое может быть оценено, как объяснено в [Thiergart2012]. В качестве альтернативы,  $H(k, n)$  может быть найден посредством оценки, так называемой диффузности, как описано в [Pulkki2007, Thiergart2013]. Оценка SDR или диффузности обычно требует более чем одного микрофона. Тем не менее, рассеянный звук, в конце концов, достигается посредством фильтрации единственного микрофонного сигнала.

Пример системы для извлечения рассеянного звука с использованием одноканального фильтра показан на фиг. 1. Во-первых, SDR (или, альтернативно, диффузность) оценивается на основании множества микрофонных сигналов. Во-вторых, вычисляется фильтр  $H(k, n)$  на основании этой информации. Наконец, фильтр  $H(k, n)$  умножается на единственный микрофонный сигнал для получения оценки рассеянного звука.

Многоканальные фильтры учитывают  $M > 1$  микрофонов. Такие фильтры использовались, например, в [Thiergart2013b, Kowalczyk2013]. Для вывода следующих формул, представим  $M$  микрофонных сигналов вектором

$\mathbf{x}(k, n) = [X_1(k, n), X_2(k, n), \dots, X_M(k, n)]^T$ . Рассеянный звук в  $m$ -ом микрофоне оценивается через линейную комбинацию  $M$  микрофонных сигналов, т.е.

$$\hat{X}_d(k, n, \mathbf{d}_m) = \mathbf{w}_m^H(k, n) \mathbf{x}(k, n)$$

где  $\mathbf{w}_m$  является комплексным весовым вектором длины  $M$ . Веса  $\mathbf{w}_m$  должны быть вычислены таким образом, чтобы была получена точная оценка рассеянного звука.

Непосредственным способом нахождения подходящего фильтра является вычисление весов  $\mathbf{w}_m$  таким образом, чтобы  $L$  плоских волн были подавлены, при минимизации стационарного шума  $X_n(k, n, \mathbf{d}_m)$ , который содержится в микрофонных сигналах. В математическом выражении, веса фильтров представляются в виде

$$\mathbf{w}_m(k, n) = \underset{\mathbf{w}}{\operatorname{argmin}} \mathbf{w}^H \Phi_n(k) \mathbf{w},$$

с учетом линейных ограничений

$$\mathbf{w}^H \mathbf{a}_l(k, n) = 0 \quad \forall l$$

Здесь,  $\Phi_n$  является PSD-матрицей (матрицей спектральной плотности мощности) стационарного шума, т.е.  $\Phi_n = E\{\mathbf{x}_n \mathbf{x}_n^H\}$ , которая может быть оценена с использованием общеизвестных методик, например, когда отсутствует рассеянный звук и прямой звук. Кроме того,  $\mathbf{a}_l$  является так называемым вектором распространения. Его элементами являются относительные передаточные функции  $l$ -ой плоской волны от  $m$ -го микрофона



к другим микрофонам. Следовательно,  $\mathbf{a}_l$  является вектором-столбцом длины  $M$  (следует понимать, что только рассеянный звук в  $m$ -ом микрофоне оценивается посредством  $w_m$  – взвешенной линейной комбинации  $M$  микрофонных сигналов; рассеянный звук в других микрофонах является, по существу, излишним, поскольку эти сигналы связаны через относительные передаточные функции от  $m$ -го микрофона к другим микрофонам и могут быть вычислены таким образом, при необходимости). Элементы  $\mathbf{a}_l$  зависят от DOA  $l$ -ой плоской волны. Это означает, что  $\mathbf{a}_l$  является функцией DOA  $l$ -ой плоской волны, т.е.  $\mathbf{a}_l = f(\mathbf{n}_l)$ . Поскольку  $\mathbf{a}_l$  зависит от прямого звука (т.е. плоских волн), то далее он упоминается как ограничение прямого звука. С использованием этого пространственного фильтра мы, по существу, создаем формирователь луча, который имеет диаграмму направленности захвата с нулями в направлениях  $L$  плоских волн. В результате, все плоские волны подавляются. К сожалению, решение вышеупомянутой задачи минимизации приводит к нулевым весам  $\mathbf{w}_m$ , поскольку мы имеем только нулевые ограничения, т.е. рассеянный звук не может быть извлечен.

Для преодоления этой проблемы и предотвращения нулевых весов фильтра, [Thiergart2013b, Kowalczyk2013] предлагает использовать тот же фильтр, но с дополнительным ограничением, представленным в виде

$$\mathbf{w}^H \mathbf{a}_0(k, n) = 1$$

где  $\mathbf{a}_0$  является вектором распространения, который соответствует конкретному DOA  $\mathbf{n}_0$ , по которому никакие плоские волны не поступают. С использованием этого ограничения, можно предотвратить нулевые веса фильтров, но, все же, нельзя захватить нежелательный прямой звук. В результате, с использованием этого фильтра, захватывается только рассеянный звук и некоторый шум, а все плоские волны затухают. В [Thiergart2013b], DOA  $\mathbf{n}_0$ , которому соответствует вектор  $\mathbf{a}_0$ , находится путем выбора направления, которое имеет наибольшее угловое расстояние до всех DOA  $\mathbf{n}_l(k, n)$  плоских волн. Например, если единственная плоская волна поступает под углом 0 градусов, то тогда  $\mathbf{n}_l(k, n)$  соответствует 180 градусам. К сожалению, DOA  $\mathbf{n}_0$  не гарантирует, что мы получаем оценку рассеянного звука с настолько малым шумом, насколько это возможно. Кроме того, полученная в результате диаграмма направленности захвата является не очень оптимальной для захвата рассеянного звука, поскольку она становится остронаправленной при более высоких частотах. Это является недостатком, если целью является захват рассеянного звука со всех направлений.

Пример полученной диаграммы направленности захвата показан на фиг. 6. Здесь, два прямых звука поступают с двух азимутальных направлений  $51^\circ$  и  $97^\circ$ . Чертеж показывает полученную диаграмму направленности захвата при частоте 2,8 кГц при использовании однородного линейного массива с 16 микрофонами с интервалами 5 см между микрофонами. Диаграмма направленности захвата имеет нули точно для  $51^\circ$  и  $97^\circ$  и высокий коэффициент усиления для  $180^\circ$ , что соответствует направлению  $\mathbf{n}_0$ .

Кроме того, диаграмма направленности захвата имеет множество других пространственных нулей или низких коэффициентов усиления почти для всех других направлений. Эта диаграмма направленности захвата не подходит для захвата рассеянного звука, который поступает со всех направлений. Следует снова отметить,

что ограничения прямого звука  $\mathbf{a}_i$  прямо связаны с DOA прямых звуков. Необходимая диаграмма направленности захвата, которая не может быть получена с использованием пространственного фильтра в этой подсекции, показана на фиг. 7. Эта диаграмма направленности захвата имеет два пространственных нуля для DOA прямых звуков, а для иных направлений является почти всенаправленной. Эта диаграмма направленности захвата обеспечена с использованием предлагаемого фильтра, который описан ниже со ссылкой на фиг. 7.

Замкнутое по форме решение для вычисления весов фильтра  $\mathbf{w}_m$  с учетом вышеупомянутых ограничений может быть найдено в [VanTrees2002]. Для вычисления пространственного фильтра, необходимо знать DOA  $L$  плоских волн, а именно, вычислить ограничения прямого звука  $\mathbf{a}_i$  и  $\mathbf{a}_0$ . Эта DOA-информация может быть определена с использованием общеизвестных узкополосных алгоритмов оценки DOA, таких как Root MUSIC или ESPRIT. Следует дополнительно отметить, что элементы  $\mathbf{a}_0$  являются обычно комплексными, и  $\mathbf{a}_0$  обычно должен быть пересчитан для каждого  $k$  и  $n$ , поскольку следует предположить, что DOA плоских волн являются сильно изменяющимися во времени. Сильная флуктуация  $\mathbf{a}_0$  может привести к слышимым шумам.

Пример системы для извлечения рассеянного звука с использованием представленного многоканального фильтра показан на фиг. 2. После преобразования микрофонных сигналов в частотно-временную область, мы оцениваем стационарный шум и DOA  $L$  плоских волн. Затем, на основании DOA-информации, достигается  $M+1$  линейных ограничений прямого звука ( $\mathbf{a}_i$  и  $\mathbf{a}_0$ ). На основе этой информации, могут быть вычислены веса фильтра. Применение этих весов к микрофонным сигналам дает необходимую оценку рассеянного звука. Из этого описания ясно, что полученный фильтр действительно зависит только от прямого звука (т.е. от DOA и соответствующих относительных передаточных функций плоских волн между микрофонами, соответственно), а не от рассеянного звука. Это означает, что фильтр не учитывает потенциально доступную информацию в отношении рассеянного звука, даже если его используют для оценки рассеянного звука.

В настоящем изобретении, мы предлагаем новый многоканальный фильтр для извлечения реверберирующего звука, который преодолевает ограничения вышеупомянутых фильтров. Предлагаемый пространственный фильтр характеризуется диаграммой направленности, которая приближается к всенаправленной диаграмме направленности, за исключением направлений поступления (direction-of-arrival - DOA) прямого звука, для которых она показывает пространственные нули. Это представляет в высшей степени необходимую способность захвата рассеянного звука со всех направлений с низкими искажениями.

Далее, мы предлагаем многоканальный фильтр для оценки рассеянного звука  $X_d(k, n, \mathbf{d}_m)$  в положении  $\mathbf{d}_m$  с использованием  $M > 1$  микрофонов. В отношении многоканальных фильтров, описанных выше, давление рассеянного звука в  $m$ -ом микрофоне оценивается посредством выполнения линейной комбинации микрофонных сигналов, т.е.

$$\hat{X}_d(k, n, \mathbf{d}_m) = \mathbf{w}_m^H(k, n) \mathbf{x}(k, n).$$

Весовой вектор  $\mathbf{w}_m$ , который предложен далее, минимизирует конкретную целевую

функцию и является линейно ограниченным, подобно многоканальным фильтрам, описанным выше.

Однако, в отличие от многоканальных фильтров, описанных выше, мы предлагаем использовать линейное ограничение, которое не зависит от прямого звука (т.е. от L плоских волн). Точнее, предлагаемое новое ограничение не является функцией DOA плоских волн или соответствующих относительных передаточных функций плоских волн между микрофонами, соответственно.

Напротив, предлагаемое новое ограничение зависит от статистической информации о рассеянном звуке, т.е. предлагаемое новое ограничение зависит от относительных передаточных функций рассеянного звука между микрофонами. Мы покажем далее, что предлагаемое новое ограничение является функцией когерентности или корреляции рассеянного звука между микрофонами. Эта когерентность соответствует средней относительной передаточной функции рассеянного звука между микрофонами.

Предлагаемый пространственный фильтр достигается минимизацией конкретной целевой функции, удовлетворяя ограничение отсутствия искажений для рассеянного звука. Это ограничение соответствует относительной передаточной функции рассеянного звука между микрофонами. В математическом выражении, фильтр вычисляется следующим образом

$$\mathbf{w}_m(k, n) = \underset{\mathbf{w}}{\operatorname{argmin}} J(\mathbf{w}),$$

с учетом линейного ограничения

$$\mathbf{w}^H \mathbf{b}_m(k, n) = 1.$$

Здесь, J является целевой функцией, подлежащей минимизации фильтром. Целевая функция может быть, например, мощностью стационарного шума на выходе фильтра, энергией помехи на выходе фильтра, или квадратичной ошибкой оцененного рассеянного звука. Примеры для J будут обеспечены в вариантах осуществления. Вектор ограничения

$\mathbf{b}_m$  можно представить в виде  $\mathbf{b}_m(k, n) = [B_{1,m}(k, n), B_{2,m}(k, n), \dots, B_{M,m}(k, n)]^T$ . m'-ый элемент

$B_{m',m}$  является относительной передаточной функцией рассеянного звука между

микрофонами m и m'. Эту относительную передаточную функцию можно представить в виде

$$B_{m',m}(k, n) = \frac{X_d(k, n, \mathbf{d}_m)}{X_d(k, n, \mathbf{d}_{m'})}.$$

Следует отметить, что m-ый элемент  $\mathbf{b}_m$  равен 1. С использованием этого ограничения, мы захватываем рассеянный звук без искажений. Фактически, пусть

$\mathbf{x}_d(k, n) = [X_d(k, n, \mathbf{d}_1), X_d(k, n, \mathbf{d}_2), \dots, X_d(k, n, \mathbf{d}_M)]^T$  является вектором, содержащим

записанный рассеянный звук. С использованием уравнений, приведенных выше, этот вектор может быть записан в следующем виде

$$\mathbf{x}_d(k, n) = \mathbf{b}_m(k, n) X_d(k, n, \mathbf{d}_m).$$

Рассеянный звук на выходе фильтра можно представить в виде выражения

$\mathbf{w}^H(k, n) \mathbf{x}_d(k, n)$ , которое идентично  $X_d(k, n, \mathbf{d}_m)$ , поскольку  $\mathbf{w}^H \mathbf{b}_m(k, n) = 1$ . Таким образом,

этот фильтр захватывает рассеянный звук без искажений. Относительные передаточные функции в  $\mathbf{b}_m$  обычно не могут быть оценены на практике, поскольку они являются, по существу, случайными, т.е. мы имеем разную реализацию передаточной функции

для каждого  $k$  и  $n$ . Таким образом,  $B_{m',m}$  вычисляется как средняя относительная передаточная функция между микрофонами  $m$  и  $m'$ , т.е.

$$B_{m',m}(k,n) = \gamma_{m',m}(k,n).$$

Эта средняя относительная передаточная функция  $\gamma_{m',m}$  соответствует так называемой пространственной когерентности рассеянного звука между микрофонами  $m$  и  $m'$ , которая определяется в качестве

$$\gamma_{m',m}(k,n) = \frac{E\{X_a(k,n,\mathbf{d}_m) X_a^*(k,n,\mathbf{d}_{m'})\}}{\sqrt{E\{|X_a(k,n,\mathbf{d}_m)|^2\} E\{|X_a(k,n,\mathbf{d}_{m'})|^2\}}},$$

где  $(.)^*$  означает комплексно сопряженную величину. Эта пространственная когерентность описывает корреляцию рассеянного звука между микрофонами  $m$  и  $m'$  в частотной области. Эта когерентность зависит от конкретного рассеянного звукового поля. Когерентность может быть измерена заранее для данного помещения. Альтернативно, когерентность известна из теории конкретных рассеянных звуковых полей [Elko2001]. Например, для сферически изотропного рассеянного звукового поля, которое часто рассматривается на практике, мы имеем

$$\gamma_{m',m}(k,n) = \text{sinc}\left(2\pi \frac{f}{c} r_{m',m}\right),$$

где  $\text{sinc}$  означает  $\text{sinc}$ -функцию,  $f$  является акустической частотой для данного частотного диапазона  $k$ , и  $c$  является скоростью звука. Кроме того,  $r_{m',m}$  является расстоянием между микрофонами  $m$  и  $m'$ . При использовании пространственной когерентности в качестве линейного ограничения  $B_{m',m}$ , которое представляет среднюю относительную передаточную функцию рассеянного звука между микрофонами, полученный фильтр является эквивалентным сумме многих линейно ограниченных пространственных фильтров, причем каждый из этих фильтров захватывает разную реализацию случайного рассеянного звука без искажений.

С использованием ограничения рассеянного звука, введенного выше, мы получаем пространственный фильтр, который захватывает рассеянный звук одинаково хорошо со всех направлений. Это является отличием от многоканальных фильтров, описанных выше, которые захватывают звук, главным образом, по одному направлению, а именно, по направлению, которому соответствует выбранный вектор распространения  $\mathbf{a}_0$ .

Следует отметить, что ограничение рассеянного звука  $\mathbf{b}_m$ , концептуально, сильно отличается от ограничений прямого звука  $\mathbf{a}_i$  и  $\mathbf{a}_0$ . Таким образом, новый фильтр, предложенный в этом разделе, является, концептуально, сильно отличающимся от многоканальных фильтров, описанных выше.

Блок-схема предложенного изобретения показана на фиг. 3. Во-первых,  $M$  микрофонные сигналы преобразуются в частотно-временную область (или другую подходящую область для обработки сигналов) с использованием банка (101) фильтров (filterbank - FB). Во-вторых, мы вычисляем вектор линейного ограничения рассеянного звука  $\mathbf{b}_m$  в блоке (102). Вектор ограничения рассеянного звука оценивается либо на основании сигнала, либо он соответствует, например, теоретической пространственной когерентности для конкретного предлагаемого рассеянного поля, как упомянуто выше. В блоке (104), мы оцениваем конкретную статистику (например, статистику шумов) на

основании микрофонных сигналов. Эта информация, которая обычно представляется в виде PSD-матрицы  $\Phi(k, n)$ , используется для генерации целевой функции J, которая должна быть минимизирована фильтром. Веса фильтра, который минимизирует целевую функцию с учетом ограничения рассеянного звука, вычисляются в блоке (103). Наконец, веса применяются к микрофонным сигналам, что обеспечивает в результате необходимую оценку рассеянного звука. Конкретные реализации настоящего изобретения представлены в нижеследующих вариантах осуществления.

#### МИНИМИЗАЦИЯ ВЫХОДНОЙ МОЩНОСТИ, УДОВЛЕТВОРЯЮЩЕЙ ОГРАНИЧЕНИЮ РАССЕЯННОГО ЗВУКА

В этом варианте осуществления, мы определяем пространственный фильтр, который минимизирует всю выходную мощность фильтра с учетом ограничения рассеянного звука. Ограничение рассеянного звука обеспечивает, то что рассеянный звук сохраняется пространственным фильтром, в то время как остальные части сигнала (нежелательный стационарный шум и плоские волны) минимизируются. Веса фильтра  $\mathbf{w}_m$  вычисляются следующим образом:

$$\mathbf{w}_m(k, n) = \underset{\mathbf{w}}{\operatorname{argmin}} \mathbf{w}^H \Phi_x(k, n) \mathbf{w},$$

с учетом линейного ограничения

$$\mathbf{w}^H \mathbf{b}_m(k, n) = 1.$$

Замкнутое по форме решение для этого фильтра приведено в [VanTrees2002] в виде

$$\mathbf{w}_m^H(k, n) = \frac{\mathbf{b}_m^H(k, n) \Phi_x^{-1}(k, n)}{\mathbf{b}_m^H(k, n) \Phi_x^{-1}(k, n) \mathbf{b}_m(k, n)}.$$

Здесь,  $\Phi_x$  является PSD-матрицей микрофонных сигналов, которая может быть вычислена следующим образом:

$$\Phi_x(k, n) = E\{\mathbf{x}(k, n) \mathbf{x}^H(k, n)\},$$

где  $\mathbf{x}(k, n)$  является вектором, содержащим микрофонные сигналы. На практике, математическое ожидание аппроксимируется, например, путем временного усреднения. Кроме того, элементы вектора ограничения  $\mathbf{b}_m(k, n) = [B_{1,m}(k, n), B_{2,m}(k, n), \dots, B_{M,m}(k, n)]^T$  соответствуют пространственной когерентности рассеянного звука между микрофонами  $m$  и  $m'$ , т.е.

$$B_{m',m}(k, n) = \gamma_{m',m}(k, n).$$

Фактически, пространственная когерентность  $B_{m',m}$  не должна зависеть от времени (а именно,  $B_{m',m}(k, n) = B_{m',m}(k)$ ), следовательно, пространственная когерентность может быть оценена заранее или может предполагать теоретическое значение. Пространственная когерентность может быть оценена либо на основании микрофонных сигналов (во время периодов, когда присутствует только рассеянный звук) с использованием

$$\gamma_{m',m}(k, n) = \frac{E\{X(k, n, \mathbf{d}_m) X^*(k, n, \mathbf{d}_{m'})\}}{\sqrt{E\{|X(k, n, \mathbf{d}_m)|^2\} E\{|X(k, n, \mathbf{d}_{m'})|^2\}}},$$

либо может быть представлена в виде априорной информации, предполагающей конкретное рассеянное звуковое поле. В последнем случае, мы используем, например,

пространственную когерентность для сферически изотропного рассеянного звукового поля, т.е.

$$\gamma_{\mathbf{m},\mathbf{m}}(\mathbf{k},n) = \text{sinc}\left(2\pi \frac{f}{c} r_{\mathbf{m},\mathbf{m}}\right).$$

Следует отметить, что sinc-функция может быть заменена другими функциями, в зависимости от предполагаемого звукового поля. Для разных рассеянных звуковых полей существуют разные функции когерентности, которые априори известны. Примеры могут быть найдены в [Elko2001].

Блок-схема этого варианта осуществления показана на фиг. 3. После преобразования микрофонных сигналов с использованием банка (101) фильтров, мы вычисляем сигнальную PSD-матрицу  $\Phi_{\mathbf{x}}$  в блоке (104) оценки статистики сигналов. Кроме того, мы вычисляем вектор линейного ограничения рассеянного звука  $\mathbf{b}_{\mathbf{m}}$  в блоке (102), либо на основании сигнала, либо с использованием априорной информации, предполагающей конкретное рассеянное звуковое поле. Веса фильтра, затем, вычисляются в блоке (103). Умножение весов на микрофонные сигналы дает необходимую оценку рассеянного звука.

Фильтр, вычисленный в этом варианте осуществления, имеет следующие преимущества по сравнению с другим пространственным фильтром (например, фильтрами, описанными в предшествующем уровне техники):

- Поскольку мы минимизируем весь выходной сигнал, фильтр обеспечивает оптимальное соотношение между затуханием  $L$  плоских волн и стационарным шумом.

- Необходимую сигнальную PSD-матрицу  $\Phi_{\mathbf{x}}$ , и, таким образом, фильтр, легко вычислить на практике.

- Не требуется никакой информации о DOA  $L$  плоских волн.

#### ЛИНЕЙНО ОГРАНИЧЕННЫЙ ФИЛЬТР С МИНИМАЛЬНОЙ ДИСПЕРСИЕЙ

Этот вариант осуществления представляет собой комбинацию новой методики и методики предшествующего уровня техники в отношении многоканальных фильтров, описанных выше в соответствии с фиг. 2. В этом варианте осуществления, мы определяем линейно ограниченный пространственный фильтр, который минимизирует стационарный шум на выходе фильтра, с учетом рассеянного ограничения и дополнительных направленных ограничений. Веса фильтра  $\mathbf{w}_{\mathbf{m}}$  вычисляются в виде

$$\mathbf{w}_{\mathbf{m}}(\mathbf{k},n) = \underset{\mathbf{w}}{\text{argmin}} \mathbf{w}^H \Phi_{\mathbf{n}}(\mathbf{k}) \mathbf{w},$$

с учетом линейных ограничений

$$\mathbf{w}^H \mathbf{b}_{\mathbf{m}}(\mathbf{k},n) = 1$$

и

$$\mathbf{w}^H \mathbf{a}_l(\mathbf{k},n) = 0 \quad \forall l$$

Ясно, что фильтр минимизирует только стационарный шум на выходе. Нежелательные плоские волны подавляются с использованием вторых линейных ограничений (как объяснено выше для многоканальных фильтров, фиг. 2). Эти дополнительные ограничения, по сравнению с фильтром, минимизирующим выходную мощность, в соответствии с фиг. 3, обеспечивают еще более сильное подавление создающих помехи плоских волн. Полученный в результате фильтр, тем не менее, сохраняет рассеянный звук благодаря первому линейному ограничению. Замкнутое по форме решение для этого фильтра, которое может быть вычислено на практике, может быть представлено



В ВИДЕ

$$\mathbf{w}_m^H(k, n) = \mathbf{g}^H(k, n) \left[ \mathbf{C}^H(k, n) \Phi_n^{-1}(k) \mathbf{C}(k, n) \right]^{-1} \mathbf{C}^H(k, n) \Phi_n^{-1}(k)$$

Здесь, вектор  $\mathbf{C} = [\mathbf{b}_m, \mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_L]$  является матрицей ограничений, содержащей

линейные ограничения, определенные выше, и  $\mathbf{g} = [\mathbf{1}, \mathbf{0}]^T$  ( $\mathbf{0}$  является нулевым вектором длины  $L$ ) являются соответствующими характеристиками. В отношении многоканального фильтра, показанного на фиг. 2, векторы  $\mathbf{a}_l$  зависят от DOA  $L$  плоских волн и могут быть вычислены общеизвестным из литературы способом [VanTrees2002].  
Напротив, элементы  $\mathbf{b}_m$  описывают корреляцию или когерентность рассеянного звука между микрофонами. Элементы  $\mathbf{b}_m$  вычисляются, как объяснено в соответствии с фиг. 3. Кроме того,  $\Phi_n$  является PSD-матрицей стационарного шума. Эта PSD-матрица может быть оценена, например, во время речевых пауз. Если стационарные шумы в разных микрофонах являются взаимно независимыми, то мы можем просто заменить  $\Phi_n$  матрицей тождественности размера  $M$  на  $M$ .

Блок-схема этого варианта осуществления показана на фиг. 4. После преобразования микрофонных сигналов с использованием банка (101) фильтров, мы вычисляем PSD-матрицу  $\Phi_n$  стационарного шума в блоке (104) оценки статистики шумов. Кроме того, мы вычисляем вектор линейного ограничения рассеянного звука  $\mathbf{b}_m$  в блоке (102), либо на основании сигнала, либо с использованием априорной информации, предполагающей конкретное рассеянное звуковое поле. В блоке (105), мы оцениваем DOA  $L$  плоских волн. На основании этой информации, мы вычисляем ограничения прямого звука  $\mathbf{a}_l$  в блоке (106). Вычисленная информация передается в блок (103) вычисления фильтра, который вычисляет веса фильтра  $\mathbf{w}_m$  с использованием замкнутого по форме решения, представленного выше. Умножение весов на микрофонные сигналы дает необходимую оценку рассеянного звука.

Пример полученной диаграммы направленности захвата для этого фильтра показан на фиг. 7. Здесь, два прямых звука поступают с двух азимутальных направлений  $51^\circ$  и  $97^\circ$ . Чертеж показывает полученную диаграмму направленности захвата при частоте 2,8 кГц при использовании однородного линейного массива с 16 микрофонами с интервалами 5 см между микрофонами. Диаграмма направленности захвата имеет нули точно для  $51^\circ$  и  $97^\circ$ . В противном случае она является почти всенаправленной. Это является главным преимуществом перед пространственным фильтром предшествующего уровня техники, показанным на фиг. 2, который создает пространственный фильтр на фиг. 6.

Фильтр, вычисленный в этом варианте осуществления, имеет следующие преимущества по сравнению с другим пространственным фильтром (например, фильтрами, описанными в предшествующем уровне техники):

- Сильное затухание плоских волн благодаря ограничениям прямого звука.
- Почти всенаправленная диаграмма направленности захвата, которая необходима для захвата рассеянного звука.

#### КОМБИНИРОВАННАЯ МЕТОДИКА

Пространственные фильтры, показанные на фиг. 3 и 4, в общем, обеспечивают хорошую производительность на практике. Однако, они также страдают от конкретных недостатков. Например, фильтр, показанный на фиг. 3, обычно не подавляет прямой

звук полностью. Оставшаяся мощность прямого звука может привести к нежелательным эффектам во время воспроизведения пространственного звука. Напротив, пространственный фильтр фиг. 4 обеспечивает сравнительно плохую устойчивость к стационарному шуму при низких частотах. Далее, мы объединяем пространственные фильтры на фиг. 3 и 4 для использования преимуществ обоих пространственных фильтров одновременно смягчая недостатки. В результате, мы получаем пространственный фильтр, который характеризуется хорошим подавлением помех (подавлением прямого звука) и высокой устойчивостью к стационарному шуму.

Методика, предлагаемая ниже, представляет собой так называемый Параметрический многоканальный фильтр Винера (parametric multi-channel Wiener filter - PMWF), который может быть масштабирован между так называемым пространственным фильтром с Минимальной среднеквадратичной погрешностью (minimum mean square error - MMSE) и пространственным фильтром на фиг. 3. Дополнительно, пространственный фильтр на фиг. 4 используется для оценки конкретных необходимых величин.

Весовой вектор предлагаемого PMWF вычисляется в виде

$$\mathbf{w}_m(k, n) = \underset{\mathbf{w}}{\operatorname{argmin}} \mathbf{w}^H \Phi_x(k) \mathbf{w},$$

с учетом

$$E \left\{ \left| X_d(k, n, \mathbf{d}_m) - \hat{X}_d(k, n, \mathbf{d}_m) \right|^2 \right\} < \sigma^2,$$

где  $\sigma^2$  является максимальной квадратичной абсолютной погрешностью оцениваемого рассеянного звука. Решение этой оптимизационной задачи приводит к

$$\mathbf{w}_m(k, n) = \frac{\varphi_d(k, n)}{\beta + 1} \Phi_x^{-1}(k, n) \mathbf{b}_m,$$

где мы определяем как

$$\beta = \alpha (\varphi_d \mathbf{b}_m^H \Phi_x^{-1} \mathbf{b}_m)$$

Здесь,  $\alpha \in [0, 1]$  является определяемым пользователем управляющим параметром.

Для  $\alpha = 0$ , мы получаем MMSE пространственный фильтр, который минимизирует среднеквадратичную погрешность оцениваемого рассеянного звука. Для  $\alpha = 1$ , мы получаем пространственный фильтр, предлагаемый на фиг. 3. Для промежуточных значений  $\alpha$ , мы имеем масштабирование между этими двумя пространственными фильтрами. Меньшее  $\alpha$  обеспечивает лучший шум и подавление помех, в то время как более высокое  $\alpha$  обеспечивает оценку рассеянного звука с меньшими искажениями.

Следует отметить, что элементы  $\mathbf{b}_m$  вычисляются, как объяснено в соответствии с фиг.

3. Мощность рассеянного звука  $\phi_d$  может быть вычислена с использованием пространственного фильтра, предлагаемого на фиг. 4, который обеспечивает очень хорошее подавление плоских волн. Пусть  $\mathbf{w}_1$  означает пространственный фильтр на фиг. 4 для оценки рассеянного звука в первом микрофоне. Учитывая  $\mathbf{w}_1$  мощность рассеянного звука в первом микрофоне может быть получена, например, как описано в [Thiergart2013b], т.е.

$$\hat{\phi}_d(k, n) = \frac{\mathbf{w}_1^H(k, n) [\Phi_x(k, n) - \Phi_n(k, n)] \mathbf{w}_1(k, n)}{\mathbf{w}_1^H(k, n) \Gamma_d \mathbf{w}_1(k, n)}$$

или с использованием методики с управлением по решению, как объяснено в



[Kowalczyk2013]. В предшествующей формуле,  $\Gamma_d$  является матрицей пространственной когерентности  $M \times M$  для рассеянного звука.  $(m, m')$ -ый элемент  $\Gamma_d$  является пространственной когерентностью  $\gamma_{m',m}$  между микрофонами  $m$  и  $m'$ . Эта

Блок-схема этого варианта осуществления показана на фиг. 5. После преобразования микрофонных сигналов с использованием банка (101) фильтров, мы вычисляем микрофонную PSD-матрицу  $\Phi_x$  и шумовую PSD-матрицу  $\Phi_n$  в блоке (104). Кроме того,

мы вычисляем линейное ограничение рассеянного звука  $\mathbf{b}_m$  в блоке (102), либо на основании сигнала, либо с использованием априорной информации, предполагающей конкретное рассеянное звуковое поле. В блоке (105), мы оцениваем DOA  $L$  плоских волн. На основании этой информации, мы вычисляем ограничения прямого звука  $\mathbf{a}_l$  в блоке (106). Эти ограничения используются в (107) вместе с  $\Phi_n$  для вычисления весов  $\mathbf{w}_l$ . Мощность рассеянного звука  $\phi_d$  вычисляется в (108) на основании  $\mathbf{w}_l$  и  $\Phi_n$ .

Конечные веса пространственного фильтра  $\mathbf{w}_m$  могут быть, затем, вычислены в (103) с использованием  $\phi_d$ ,  $\Phi_x$ , и  $\mathbf{b}_m$ . С использованием параметра  $\alpha$ , можно масштабировать пространственный фильтр между MMSE-фильтром и PMWF. Умножение весов  $\mathbf{w}_m$  на микрофонные сигналы дает необходимую оценку рассеянного звука.

Фиг. 8 схематично показывает микрофонный массив, содержащий  $M$  микрофонов. Микрофонный массив подвергается воздействию звукового поля, содержащего участки прямого звука и участки рассеянного звука. Участки прямого звука представлены на фиг. 8 посредством  $L$  источников звука от  $X_{l=1}$  до  $X_{l=L}$ , каждый из которых, как предполагается, генерирует, по существу, плоскую волну в месте расположения микрофонного массива. Направление распространения каждой плоской волны (в месте расположения микрофонного массива) указано на фиг. 8 посредством векторов  $\mathbf{n}_1$ - $\mathbf{n}_L$ . Участки прямого звука являются обычно функцией местоположения  $\mathbf{d}_m$ . Рассеянный звук  $X_d(k, n, \mathbf{d}_m)$  может моделировать, например, реверберацию в помещении. Рассеянный звук, как предполагается, генерируется посредством бесконечной суммы плоских волн со случайными фазами, случайной амплитудой, и случайными DOA. Это означает, что рассеянный звук генерируется посредством бесконечного количества источников, случайно распределенных по звуковой сцене. Эти источники звука моделируют бесконечное количество отражений помещения, которые генерируют позднюю реверберацию.

Отдельный микрофонный сигнал, например, для  $m$ -ого микрофона,  $X_m(k, n)$ , является комбинацией  $L$  участков прямого звука от  $X_{l=1}$  до  $X_{l=L}$ , участка рассеянного звука  $X_d$ , и шума  $X_n$ , т.е.

$$X_m(k, n) = \sum_{l=1}^L X_l(k, n, \mathbf{d}_m) + X_d(k, n, \mathbf{d}_m) + X_n(k, n, \mathbf{d}_m).$$

Относительные передаточные функции  $B_{1,m}$ ,  $B_{2,m}$ , ...,  $B_{m',m}$ , ...,  $B_{M,m}$  для рассеянного звука между другими микрофонами и  $m$ -ым микрофоном схематично показаны на фиг. 8. Относительная передаточная функция  $B_{m,m}$  от  $m$ -ого микрофона к самому себе (не показана на фиг. 8) является обычно равной 1.

Нижеследующий список обеспечивает краткий обзор некоторых аспектов, которые были описаны выше:

- Прием по меньшей мере двух микрофонных сигналов.
- Преобразование микрофонных сигналов в частотно-временную область или другую 5 подходящую область.
- Вычисление линейного ограничения рассеянного звука в виде функции корреляции или когерентности рассеянного звука между микрофонами.
- Вычисление статистики сигналов/шумов.
- В некоторых вариантах осуществления: оценка DOA прямых звуков и вычисление 10 ограничений прямого звука, которые представляют собой относительные передаточные функции прямых звуков между микрофонами.
- В некоторых вариантах осуществления: вычисление вспомогательного фильтра и оценка мощности рассеянного звука.
- Вычисление весов пространственного фильтра для извлечения рассеянного звука 15 с использованием полученной статистики сигналов/шумов и необязательной информации о мощности рассеянного звука, с учетом ограничения рассеянного звука.
- Выполнение линейной комбинации микрофонных сигналов с использованием вычисленных весов пространственного фильтра.

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

Аналогично, аспекты, описанные в контексте способа, также представляют собой описание соответствующего блока или элемента или признака соответствующего 25 устройства.

В зависимости от конкретных требований к реализации, варианты осуществления 30 настоящего изобретения могут быть реализованы в аппаратном обеспечении или программном обеспечении. Реализация может быть выполнена с использованием цифровой запоминающей среды, например, гибкого диска, DVD, CD, ROM, PROM, EPROM, EEPROM или FLASH-памяти, на которой сохранены электронным способом считываемые управляющие сигналы, которые взаимодействуют (или способны 35 взаимодействовать) с программируемой вычислительной системой таким образом, чтобы соответствующий способ выполнялся.

Некоторые варианты осуществления согласно настоящему изобретению содержат 40 постоянный носитель данных, на котором находятся считываемые электронными средствами управляющие сигналы, которые способны взаимодействовать с программируемой вычислительной системой таким образом, чтобы один из способов, описанных здесь, выполнялся.

В общем, варианты осуществления настоящего изобретения могут быть реализованы 45 в виде компьютерного программного продукта с использованием программного кода, причем программный код действует для выполнения одного из способов, когда компьютерный программный продукт выполняется на компьютере. Программный код может, например, храниться на машиночитаемом носителе.

Другие варианты осуществления содержат компьютерную программу для выполнения 50 одного или более способов, описанных в материалах настоящей заявки, хранящуюся на машиночитаемом носителе.

Другими словами, вариант осуществления способа согласно изобретению является, таким образом, компьютерной программой, имеющей программный код для выполнения 55 одного из способов, описанных в материалах настоящей заявки, когда компьютерная

программа выполняется на компьютере.

Дополнительный вариант осуществления способов согласно изобретению является, таким образом, носителем данных (или цифровым запоминающим носителем, или машиночитаемым носителем), содержащим записанную на нем компьютерную

5 программу для выполнения одного из способов, описанных в материалах настоящей заявки.

Дополнительный вариант осуществления способов согласно изобретению является, таким образом, потоком данных или последовательностью сигналов, представляющими компьютерную программу для выполнения одного из способов, описанных в материалах

10 настоящей заявки. Поток данных или последовательность сигналов могут быть, например, выполнены с возможностью их переноса через соединение для передачи данных, например, через интернет.

Дополнительный вариант осуществления содержит средство обработки, например, компьютер или программируемое логическое устройство, выполненное с возможностью

15 или адаптированное для выполнения одного из способов, описанных в материалах настоящей заявки.

Дополнительный вариант осуществления содержит компьютер, на котором установлена компьютерная программа для выполнения одного из способов, описанных в материалах настоящей заявки.

В некоторых вариантах осуществления может быть использовано программируемое логическое устройство (например, массив программируемых логических вентилей) для

20 выполнения некоторых или всех функциональностей способов, описанных в материалах настоящей заявки. В некоторых вариантах осуществления, массив программируемых логических вентилей может взаимодействовать с микропроцессором для выполнения

25 одного или более способов, описанных в материалах настоящей заявки. В общем, данные способы предпочтительно выполняются любым устройством, относящимся к аппаратному обеспечению.

Описанные выше варианты осуществления являются лишь иллюстрацией принципов настоящего изобретения. Следует понимать, что модификации и варианты схем и

30 элементов, описанных в материалах настоящей заявки, будут понятны специалистам в данной области техники. Таким образом, предполагается, что объем данного изобретения ограничен только объемом нижеследующих пунктов формулы изобретения, а не конкретными подробностями, представленными для описания и объяснения вариантов осуществления в материалах настоящей заявки.

### СПИСОК ЛИТЕРАТУРЫ

[Pulkki2007] - V. Pulkki, «Spatial sound reproduction with directional audio coding», J. Audio Eng. Soc, Том 55, № 6, страницы 503-516, июнь 2007.

[Thiergart2013] - O. Thiergart, G. Del Galdo, M. Taseska, and E. Habets, «Geometry-based spatial sound acquisition using distributed microphone arrays», Audio, Speech, and Language

40 Processing, IEEE Transactions on, Том 21, № 12, страницы 2583-2594, декабрь 2013.

[Kowalczyk2013] - K. Kowalczyk, O. Thiergart, A. Craciun, and E. A. P. Habets, «Sound acquisition in noisy and reverberant environments using virtual», in Applications of Signal Processing to Audio and Acoustics (WASPAA), 2013 IEEE Workshop on, октябрь 2013.

[Thiergart2013b] - 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, страницы 659-663.

[Thiergart2012] - 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, Том 132, № 4, страницы. 2337– 2346, 2012.

[VanTrees2002] - H. L. Van Trees, Detection, Estimation, and Modulation Theory: Part IV: Array Processing. John Wiley & Sons, апрель 2002, том 1.

[Elko2001] - G. W. Elko, «Spatial coherence functions for differential micro- phones in isotropic noise fields», in Microphone Arrays: Signal Processing Techniques and Applications, под редакцией M. Brandstein и D. Ward, глава 4, 61–85 (Springer, Berlin) (2001).

#### (57) Формула изобретения

1. Способ получения коэффициентов фильтрации для фильтра рассеянного звука, причём способ содержит этапы, на которых:

определяют линейное ограничение для коэффициентов фильтрации фильтра рассеянного звука, причем линейное ограничение основано на пространственной когерентности между первым участком рассеянного звука в первом микрофонном сигнале и вторым участком рассеянного звука во втором микрофонном сигнале, причем первый микрофонный сигнал захвачен первым микрофоном, а второй микрофонный сигнал захвачен вторым микрофоном, расположенным на расстоянии от первого микрофона, известным способом;

вычисляют по меньшей мере одно из направления поступления по меньшей мере одного прямого звука, статистику сигналов в отношении первого и второго микрофонных сигналов и статистику шумов в отношении первого и второго микрофонных сигналов; и

определяют коэффициенты фильтрации фильтра рассеянного звука путем решения оптимизационной задачи в отношении по меньшей мере одного из направления поступления по меньшей мере одного прямого звука, статистики сигналов и статистики шумов, с учетом линейного ограничения для коэффициентов фильтрации.

2. Способ по п. 1, дополнительно содержащий этап, на котором обеспечивают пространственную когерентность на основании относительной передаточной функции или корреляции рассеянного звука между первым микрофоном и вторым микрофоном.

3. Способ по п. 1, в котором пространственная когерентность основана на предшествующем измерении относительной передаточной функции или корреляции рассеянного звука для данной среды в течение периодов времени, когда никакой прямой звук не присутствует в среде.

4. Способ по п. 1, в котором пространственная когерентность основана на теоретическом соотношении для рассеянного звука, причем соответствующее предполагаемое рассеянное звуковое поле предполагает теоретические характеристики, относящиеся к корреляции рассеянного звука между первым микрофоном и вторым микрофоном.

5. Способ по п. 1, в котором оптимизационная задача выражена посредством

$$\mathbf{w}_m(k, n) = \underset{\mathbf{w}}{\operatorname{argmin}} J(\mathbf{w}),$$

с учетом линейного ограничения

$$\mathbf{w}^H \mathbf{b}_m(k, n) = 1,$$

где

$\mathbf{w}(k, n)$  является вектором коэффициентов фильтрации фильтра рассеянного звука;  $\mathbf{w}_m(k, n)$  является решением оптимизационной задачи на основе оценки микрофонного сигнала в m-ом микрофоне;

$J(\mathbf{w})$  является целевой функцией;

$\mathbf{b}_m(k,n)$  является вектором оцененных пространственных когерентностей, причем  $m'$ -ый элемент этого вектора является оцененной пространственной когерентностью рассеянного звука между  $m$ -ым микрофоном и  $m'$ -ым микрофоном;

$k$  является индексом частотной области; и

$n$  является индексом временной области.

6. Способ по п. 5, в котором целевая функция  $J(\mathbf{w})$  основана на одном из статистики шумов, матрицы спектральной плотности мощности (PSD) шумов, статистики сигналов, или матрицы спектральной плотности мощности (PSD) микрофона.

7. Способ по п. 1, дополнительно содержащий этапы, на которых:

оценивают по меньшей мере одно из направления поступления по меньшей мере одного прямого звука или относительной передаточной функции по меньшей мере одного прямого звука между первым микрофоном и вторым микрофоном;

вычисляют по меньшей мере одно ограничение прямого звука с использованием направления поступления или относительной передаточной функции по меньшей мере одного прямого звука, причем по меньшей мере одно ограничение прямого звука приводит к подавлению по меньшей мере одного прямого звука.

8. Способ по п. 1, в котором решением оптимизационной задачи является

$$\mathbf{w}_m(k,n) = \frac{\phi_d(k,n)}{\beta + 1} \Phi_x^{-1}(k,n) \mathbf{b}_m,$$

где

$$\beta = \alpha (\phi_d \mathbf{b}_m^H \Phi_x^{-1} \mathbf{b}_m),$$

причем

$\mathbf{w}_m(k,n)$  является решением оптимизационной задачи на основании оценки микрофонного сигнала в  $m$ -ом микрофоне;

$\mathbf{b}_m(k,n)$  является вектором оцененных пространственных когерентностей, причем  $m'$ -ый элемент вектора является оцененной пространственной когерентностью рассеянного звука между  $m$ -ым микрофоном и  $m'$ -ым микрофоном;

$\alpha \in [0,1]$  является определяемым пользователем управляющим параметром, посредством которого фильтр рассеянного звука может масштабироваться между пространственным фильтром с минимальной среднеквадратичной погрешностью и фильтром, который минимизирует выходную мощность, удовлетворяя ограничению рассеянного звука;

$\phi_d$  является мощностью рассеянного звука;

$\Phi_x$  является матрицей спектральной плотности мощности микрофонных сигналов.

9. Способ по п. 8, дополнительно содержащий этап, на котором

оценивают мощность рассеянного звука  $\phi_d$  на основании вспомогательного фильтра рассеянного звука.

10. Способ по п. 9, в котором оценку мощности рассеянного звука  $\phi_d$  выполняют на основании

$$\hat{\phi}_d(k,n) = \frac{\mathbf{w}_1^H(k,n) [\Phi_x(k,n) - \Phi_n(k,n)] \mathbf{w}_1(k,n)}{\mathbf{w}_1^H(k,n) \Gamma_d \mathbf{w}_1(k,n)},$$

причем

$\hat{\phi}_d$  является оцененной мощностью рассеянного звука, которая обеспечивает оценку

для мощности рассеянного звука  $\phi_d$ ;

$w_1$  является вектором коэффициентов фильтрации вспомогательного фильтра рассеянного звука;

5  $\Phi_x$  является матрицей спектральной плотности мощности микрофонных сигналов;

$\Phi_n$  является матрицей спектральной плотности мощности шума в пределах микрофонных сигналов; и

$\Gamma_d$  является матрицей пространственной когерентности для рассеянного звука,  
10 причем  $(m, m')$ -ый элемент  $\Gamma_d$  является пространственной когерентностью  $\gamma_{m', m}$  между микрофонами  $m$  и  $m'$ .

11. Способ по п. 1, дополнительно содержащий этап, на котором выполняют линейную комбинацию первого микрофонного сигнала и второго микрофонного сигнала с использованием коэффициентов фильтрации фильтра  
15 рассеянного звука.

12. Машиночитаемый носитель, на котором сохранена компьютерная программа для реализации способа по п. 1 при ее исполнении на компьютере или в процессоре обработки сигналов.

13. Устройство для получения коэффициентов фильтрации для фильтра рассеянного  
20 звука, причём устройство содержит:

вычислитель линейного ограничения, выполненный с возможностью определения линейного ограничения для коэффициентов фильтрации фильтра рассеянного звука, причём линейное ограничение основано на пространственной когерентности между  
25 первым участком рассеянного звука в первом микрофонном сигнале и вторым участком рассеянного звука во втором микрофонном сигнале, причём первый микрофонный сигнал захвачен первым микрофоном, а второй микрофонный сигнал захвачен вторым микрофоном, расположенным на расстоянии от первого микрофона, известным способом;

30 вычислитель, выполненный с возможностью вычисления по меньшей мере одного из направления поступления по меньшей мере одного прямого звука, статистики сигналов в отношении первого и второго микрофонных сигналов и статистики шумов в отношении первого и второго микрофонных сигналов; и

35 вычислитель коэффициентов фильтрации, выполненный с возможностью определения коэффициентов фильтрации фильтра рассеянного звука, путем решения оптимизационной задачи в отношении по меньшей мере одного из направления поступления по меньшей мере одного прямого звука, статистики сигналов и статистики шумов, с учетом линейного ограничения для коэффициентов фильтрации.

14. Устройство по п. 13, в котором пространственная когерентность основана на относительной передаточной функции рассеянного звука между первым микрофоном  
40 и вторым микрофоном.

15. Устройство по п. 13, дополнительно содержащее:

45 вспомогательный вычислитель коэффициентов фильтрации, выполненный с возможностью определения вспомогательных коэффициентов фильтрации вспомогательного фильтра рассеянного звука, путем решения другой оптимизационной задачи, с учетом другого линейного ограничения для вспомогательных коэффициентов фильтрации, отличного от линейного ограничения, учитываемого вычислителем коэффициентов фильтрации;

вспомогательный фильтр рассеянного звука, выполненный с возможностью оценки

оцененной мощности рассеянного звука;

причем вычислитель коэффициентов фильтрации выполнен с возможностью учета оцененной мощности рассеянного звука при определении коэффициентов фильтрации фильтра рассеянного звука.

5

10

15

20

25

30

35

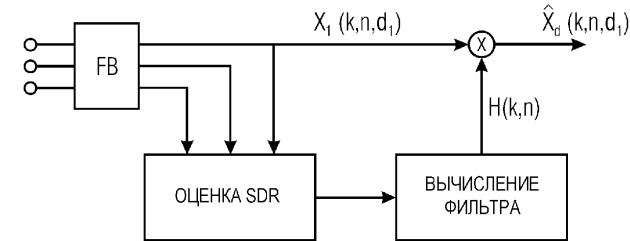
40

45

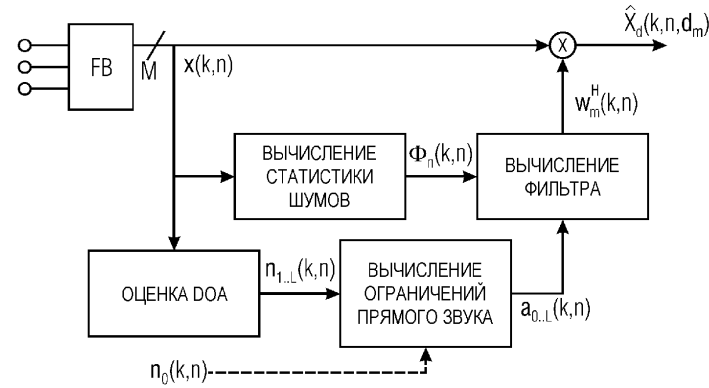
1

534978

1/7



ФИГ. 1

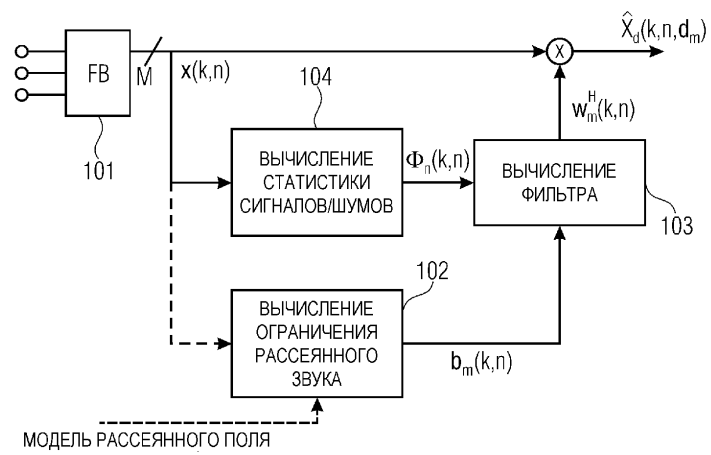


ФИГ. 2

2

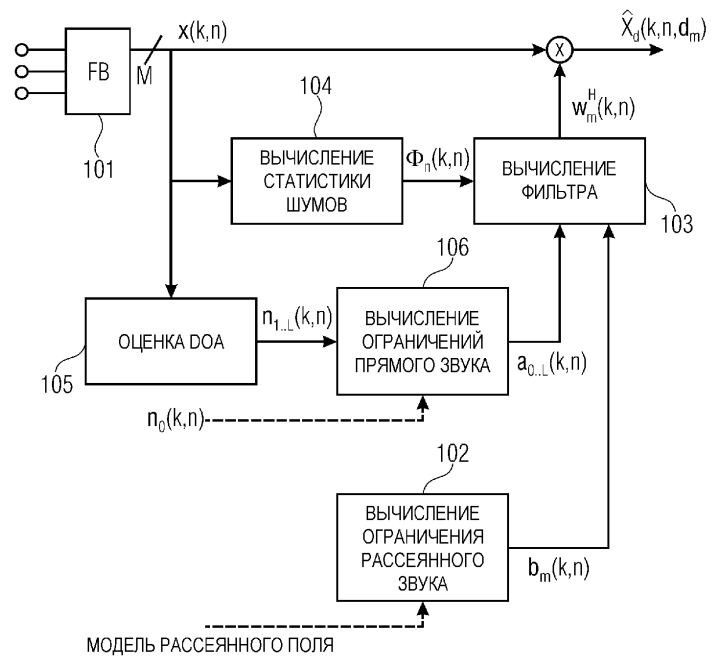


2/7

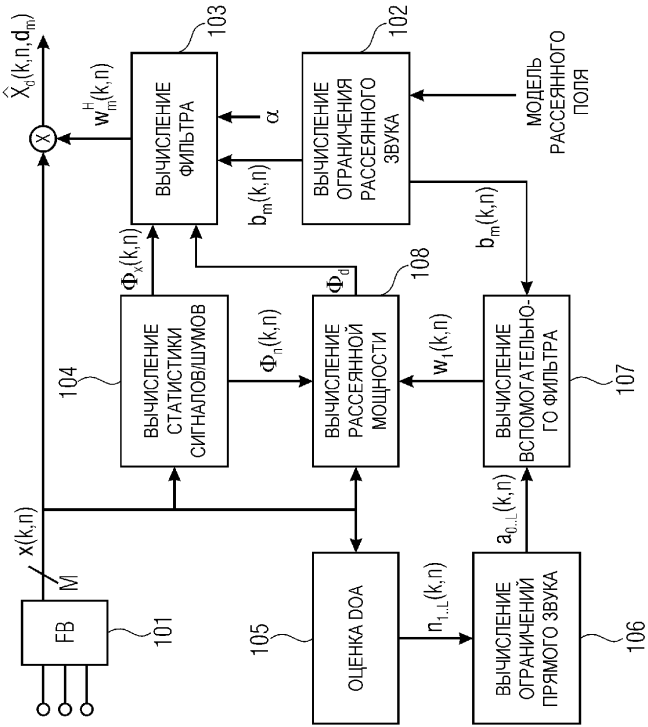


ФИГ. 3

3/7

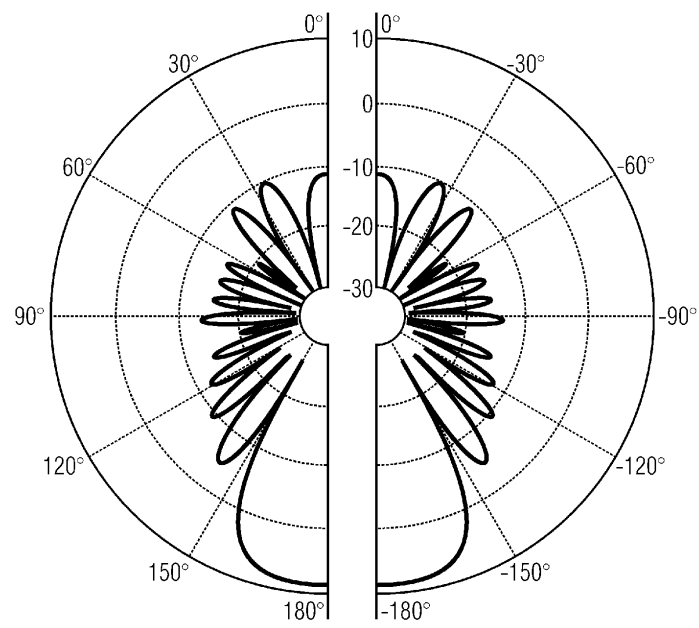


ФИГ. 4



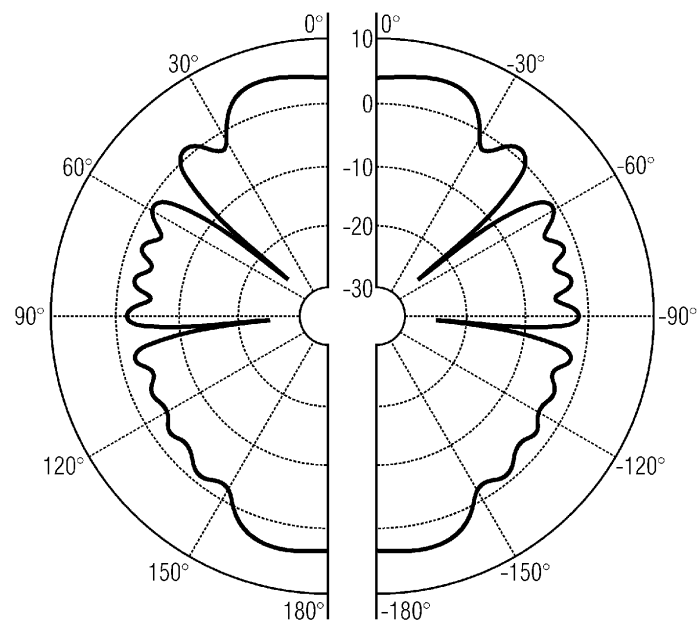
ФИГ. 5

5/7

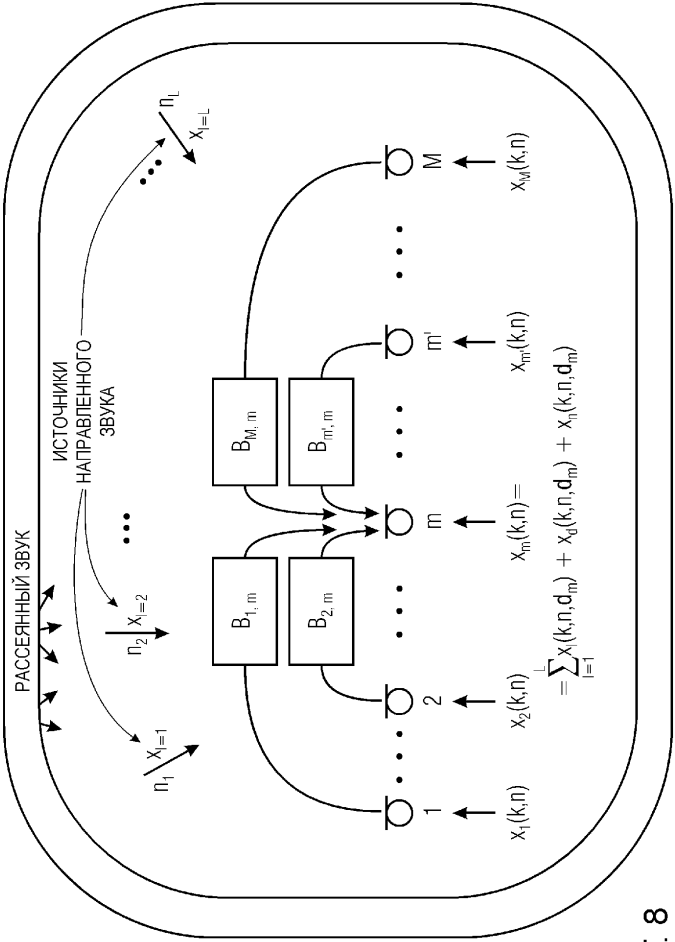


ФИГ. 6

6/7



ФИГ. 7



ФИГ. 8