



**ФЕДЕРАЛЬНАЯ СЛУЖБА
ПО ИНТЕЛЛЕКТУАЛЬНОЙ СОБСТВЕННОСТИ**

(12) ОПИСАНИЕ ИЗОБРЕТЕНИЯ К ПАТЕНТУ

(21)(22) Заявка: 2013144058/08, 24.02.2012

(24) Дата начала отсчета срока действия патента:
24.02.2012

Приоритет(ы):

(30) Конвенционный приоритет:
02.03.2011 US 61/448,444;
27.06.2011 EP 11171488.7

(43) Дата публикации заявки: 10.04.2015 Бюл. № 10

(45) Опубликовано: 10.05.2015 Бюл. № 13

(56) Список документов, цитированных в отчете о поиске: WO 2009/039897 A1, 02.04.2009. RU 2330390 C2, 27.07.2008. EP2154911 A1, 17.02.2010. EP 1565036 A2, 17.08.2005

(85) Дата начала рассмотрения заявки РСТ на национальной фазе: 02.10.2013

(86) Заявка РСТ:
EP 2012/053193 (24.02.2012)

(87) Публикация заявки РСТ:
WO 2012/116934 (07.09.2012)

Адрес для переписки:

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

(72) Автор(ы):

УЛЕ Кристиан (DE),
ПАУЛУС Йоуни (DE),
ХЕРРЕ Юрген (DE),
ПРОКАЙН Петер (DE),
ХЕЛЛЬМУТ Оливер (DE)

(73) Патентообладатель(и):

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

(54) УСТРОЙСТВО И СПОСОБ ДЛЯ ОПРЕДЕЛЕНИЯ ПОКАЗАТЕЛЯ ДЛЯ ВОСПРИНИМАЕМОГО УРОВНЯ РЕВЕРБАЦИИ, АУДИО ПРОЦЕССОР И СПОСОБ ДЛЯ ОБРАБОТКИ СИГНАЛА

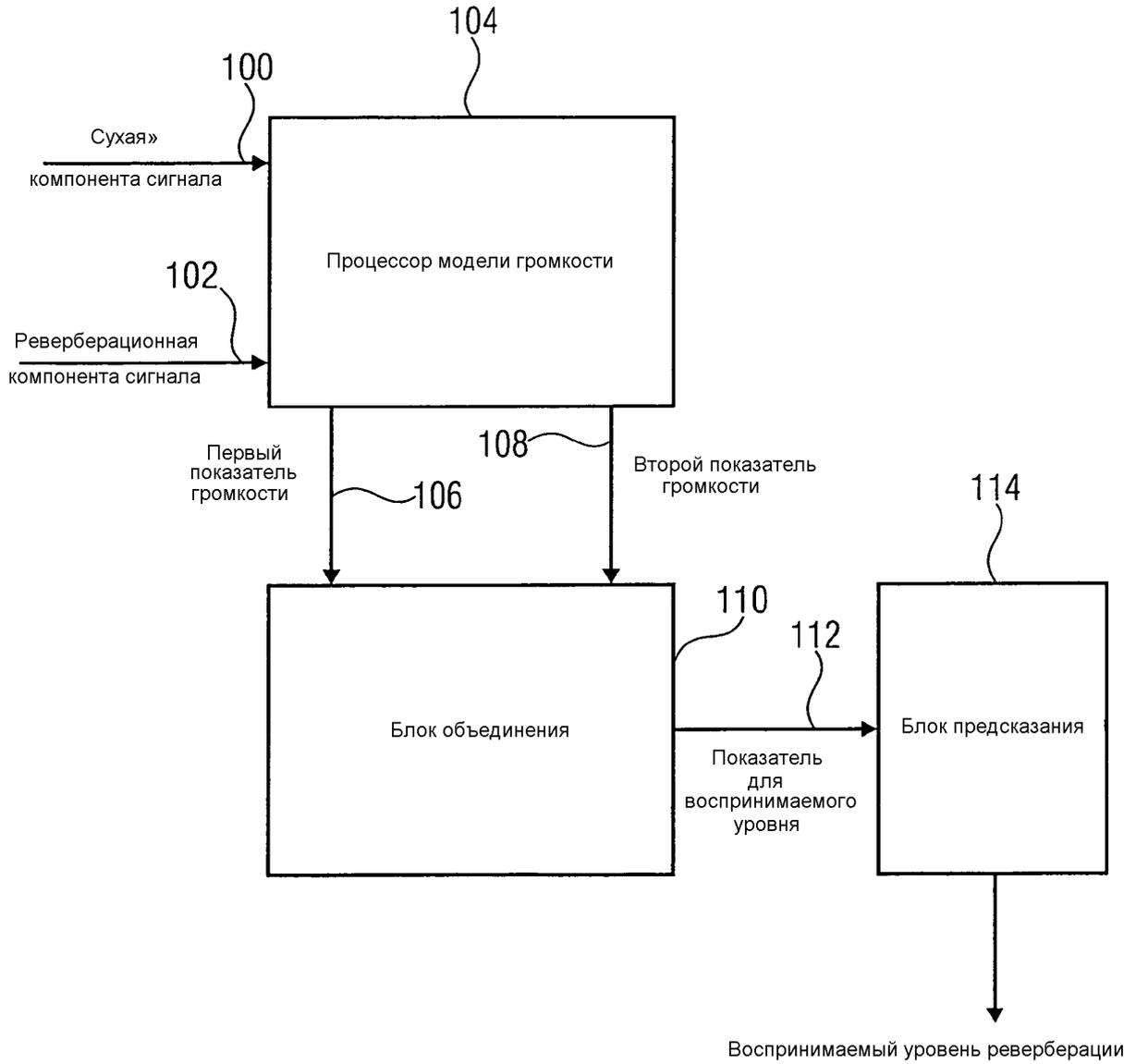
(57) Реферат:

Изобретение относится к обработке аудиосигналов и может быть применено в искусственных ревербераторах. Технический результат - улучшение характеристик воспринимаемого звука. Для этого устройство для определения показателя воспринимаемого уровня реверберации в смешанном сигнале, состоящем из прямой компоненты сигнала и реверберационной компоненты сигнала, содержит процессор модели громкости, содержащий каскад перцепционного фильтра для фильтрации «сухой» компоненты сигнала, реверберационной

компоненты сигнала или смешанного сигнала, причем каскад перцепционного фильтра выполнен с возможностью моделирования механизма слухового восприятия объекта для получения фильтрованного прямого сигнала, фильтрованного реверберационного сигнала или фильтрованного смешанного сигнала, блок оценки громкости для оценки первого показателя громкости с использованием фильтрованного прямого сигнала и для оценки второго показателя громкости с использованием фильтрованного реверберационного сигнала или фильтрованного

смешанного сигнала, где фильтрованный смешанный сигнал получен из суперпозиции прямой компоненты сигнала и реверберационной

компоненты сигнала, блок объединения для объединения первого и второго показателей громкости. 6 н. и 10 з.п. ф-лы, 17 ил., 3 табл.



ФИГ. 1

RU 2550528 C2

RU 2550528 C2



FEDERAL SERVICE
FOR INTELLECTUAL PROPERTY

(12) **ABSTRACT OF INVENTION**(21)(22) Application: **2013144058/08, 24.02.2012**(24) Effective date for property rights:
24.02.2012

Priority:

(30) Convention priority:
02.03.2011 US 61/448,444;
27.06.2011 EP 11171488.7(43) Application published: **10.04.2015** Bull. № 10(45) Date of publication: **10.05.2015** Bull. № 13(85) Commencement of national phase: **02.10.2013**(86) PCT application:
EP 2012/053193 (24.02.2012)(87) PCT publication:
WO 2012/116934 (07.09.2012)

Mail address:

129090, Moskva, ul. B. Spasskaja, 25, stroenie 3,
OOO "Juridicheskaja firma Gorodisskij i Partnery"

(72) Inventor(s):

ULE Kristian (DE),
PAULUS Jouni (DE),
KhERRE Jurgen (DE),
PROKAJN Peter (DE),
KhELL'MUT Oliver (DE)

(73) Proprietor(s):

FRAUNKhOFER-GEZELL'ShAFT T'sUR
FERDERUNG DER ANGEVANDTEN
FORShUNG E.F. (DE)(54) **DEVICE AND METHOD OF DETERMINING INDICATOR FOR PERCEPTIBLE REVERBERATION LEVEL, AUDIO PROCESSOR AND SIGNAL PROCESSING METHOD**

(57) Abstract:

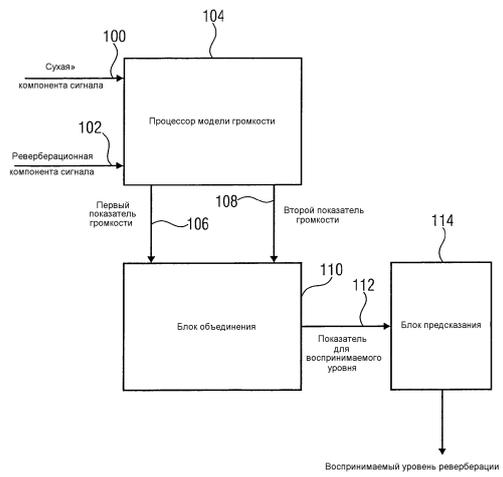
FIELD: physics.

SUBSTANCE: device for determining an indicator of a perceptible reverberation level in a composite signal, consisting of a forward signal component and a reverberation signal component, comprises a volume model processor, which comprises a perceptual filter stage for filtering the "dry" signal component, the reverberation signal component or the composite signal, wherein the perceptual filter stage is configured to simulate the audio perception mechanism of the object to obtain a filtered forward signal, a filtered reverberation signal or a filtered composite signal, a volume estimating unit for estimating a first volume indicator using the filtered forward signal and for estimating a second volume indicator using the filtered reverberation signal or the filtered composite signal, where the filtered composite signal is obtained by superposition of the forward signal component and the

reverberation signal component, an integration unit for integrating the first and second volume indicators.

EFFECT: improved characteristics of perceptible sound.

16 cl, 17 dwg, 3 tbl



ФИГ.1

Настоящая заявка относится к обработке аудиосигналов и, в частности, к обработке аудиосигналов, применимой в искусственных ревербераторах.

5 Определенный показатель для воспринимаемого уровня реверберации, например, желаемого для применений, где работает процессор искусственной реверберации в автоматическом режиме и нуждается в адаптации его параметров к входному сигналу так, чтобы воспринимаемый уровень реверберации соответствовал целевому значению. Следует отметить, что пока термин реверберация, упоминаемый в этой же теме, по-видимому, не имеет общепринятого определения, что делает его трудным для
10 использования в качестве количественного показателя теста прослушивания и сценария предсказания.

Процессоры искусственной реверберации часто реализованы как линейные инвариантные по времени системы и работают в пути отправки - возвращения сигнала, как показано на фиг.6, с предварительной задержкой d , импульсной характеристикой реверберации (ИХР) и коэффициентом масштабирования g для управления отношением
15 прямого сигнала к реверберационному сигналу (ОПР). При реализации, в качестве параметрических процессоров реверберации, они характеризуются различными параметрами, например, для управления формой и плотностью ИХР, и межканальной когерентностью (МКК) ИХР для многоканальных процессоров в одной или нескольких полосах частот.

20 Фиг.6 показывает прямой сигнал $x[k]$, входящий на входе 600, и этот сигнал передается на сумматор 602 для добавления этого сигнала к реверберационной компоненте $r[k]$ сигнала, выходящей из блока весовой обработки 604, который получает на своем первом входе сигнал, выведенный фильтром 606 реверберации, и который получает на своем втором входе, коэффициент g усиления. Фильтр 606 реверберации может иметь
25 дополнительный каскад 608 задержки, подключенный перед фильтром 606 реверберации, но из-за того, что фильтр 606 реверберации будет включать в себя некоторую задержку сам по себе, задержка в блоке 608 может быть включена в фильтр 606 реверберации так, что верхняя ветвь на фиг.6, может включать в себя только единственный фильтр, включающий в себя задержку и реверберацию, или только включающий в себя
30 реверберацию без какой-либо дополнительной задержки. Реверберационная компонента сигнала выводится фильтром 606, и эта реверберационная компонента сигнала может быть изменена множителем 606 в ответ на коэффициент g усиления для того, чтобы получить обработанную реверберационную компоненту $r[k]$ сигнала, которую затем объединяют с прямой компонентой сигнала, введенной на 600 для того, чтобы
35 окончательно получить смешанный сигнал $m[k]$ на выходе сумматора 602. Следует отметить, что термин «фильтр реверберации» относится к общим реализациям искусственных ревербераций (или в виде свертки, которая эквивалентна КИХ-фильтрации, или в виде реализаций, использующих рекурсивные структуры, такие как сети задержки обратной связи или сети всепропускающих фильтров и гребенчатые
40 фильтры обратной связи или другие рекурсивные фильтры), но означает общую обработку, которая производит реверберирующий сигнал. Такие обработки могут включать в себя нелинейные процессы или изменяющиеся во времени процессы, такие как низкочастотные модуляции амплитуд сигналов или продолжительности задержки. В этих случаях термин «фильтр реверберации» не будет применяться в строгом
45 техническом смысле линейной инвариантной по времени (ЛИВ) системы. Фактически, «фильтр реверберации» относится к обработке, которая выдает реверберирующий сигнал, возможно, включая в себя механизм для считывания рассчитанного или записанного реверберирующего сигнала из памяти.

Эти параметры оказывают влияние на полученный аудиосигнал, исходя из воспринимаемого уровня, расстояния, размера помещения, окраски и качества звука. Кроме того, воспринимаемые характеристики реверберации зависят от временных и спектральных характеристик входного сигнала [1]. Сосредоточимся на очень важном ощущении, а именно громкости, может наблюдаться, что громкость воспринимаемой реверберации монотонно связана с нестационарностью входного сигнала. Интуитивно говоря, аудиосигнал при больших колебаниях в своей огибающей возбуждает реверберацию на высоких уровнях и позволяет ей стать слышимее на более низких уровнях. В типичном сценарии, где долгосрочное ОНР, выраженное в децибелах, является положительным, прямой сигнал может маскировать реверберационный сигнал почти полностью в моменты времени, когда его энергетическая огибающая увеличивается. С другой стороны, когда сигнал прекращается, ранее возбужденный хвост реверберации проявляется в интервалах, превышающих минимальную продолжительность, определенную наклоном последующей маскировки (на максимум 200 мс) и временем интегрирования слуховой системы (при максимуме 200 мс для средних уровней).

Чтобы проиллюстрировать это, фиг.4а показывает огибающие сигнала-времени синтезированного аудиосигнала и искусственно сгенерированного реверберационного сигнала, и фиг.4б показывает функции предсказанной громкости и частичной громкости, вычисленные с вычислительной моделью громкости. Здесь используется ИХР с короткой предварительной задержкой в 50 мс, за исключением ранних отражений и синтеза поздней части реверберации с экспоненциальным затуханием белого шума [2]. Входной сигнал был сформирован из гармонического широкополосного сигнала и огибающей функции так, что воспринимаются одна волна с коротким затуханием и вторая волна с длинным затуханием. Хотя длинная волна производит больше общей энергии реверберации, становится не удивительно, что это короткий звук, который воспринимается как более реверберирующий. Где затухающий наклон более длинной волны маскирует реверберацию, короткий звук уже исчез перед созданной реверберацией и тем самым открывается интервал, в котором воспринимается реверберация. Пожалуйста, обратите внимание, что определение маскировки, использованное здесь, включает в себя как полные, так и частичные маскировки [3].

Хотя такие наблюдения были сделаны много раз [4, 5, 6], по-прежнему стоит подчеркнуть их, потому что они качественно иллюстрируют, почему модели частичной громкости могут применяться в контексте этой работы. Фактически, было указано на то, что восприятие реверберации возникает из поточных сегрегационных процессов в слуховой системе [4, 5, 6] и находится под влиянием частичного маскирования реверберации за счет прямого звука.

Приведенные выше соображения мотивируют использование моделей громкости. Похожие исследования проводились Lee et al. и фокусировались на предсказании субъективного снижения скорости ИХР при их непосредственном прослушивании [7] и на влиянии уровня воспроизведения на реверберацию [8]. Блок предсказания для реверберации, использующий основанные на громкости ранние времена затухания, предложен в [9]. В отличие от этой работы, способы предсказания, предлагаемые здесь, обрабатывают прямой сигнал и реверберационный сигнал с вычислительной моделью частичной громкости (и с ее упрощенными вариантами в поисках реализаций низкой сложности) и тем самым учитывают влияние входного (прямого) сигнала на восприятие. Недавно Tsilfidis и Mourjoroulus [10] исследовали использование модели громкости для подавления поздней реверберации в одноканальных записях. Оценка прямого сигнала

вычислена из реверберирующего входного сигнала с использованием метода спектрального вычитания, и индекс маскировки реверберации извлекается посредством вычислительной модели слухового маскирования, которая управляет обработкой реверберации.

5 Это характеристика многоканальных синтезаторов и других устройств для добавления реверберации для того, чтобы сделать звук лучше с точки зрения восприятия. С другой стороны, сгенерированная реверберация представляет собой искусственный сигнал, который при добавлении к сигналу на низком уровне, едва слышен и при
10 добавлении на высоком уровне приводит к неестественному и неприятному звучанию конечного смешанного сигнала. Что делает вещи еще хуже того, как обсуждалось в контексте фиг.4а и 4b, что воспринимаемый уровень реверберации сильно зависит от сигнала и, следовательно, некий фильтр реверберации может работать очень хорошо для одного вида сигналов, но может не иметь звукового эффекта или, что еще хуже, может генерировать серьезные звуковые артефакты для других видов сигналов.

15 Еще одна проблема связана с реверберацией в том, что реверберируемый сигнал предназначен для слуха объекта или индивидуума, такого как человек, и конечной целью генерации смешанного сигнала, имеющего прямую компоненту сигнала и реверберационную компоненту сигнала, является то, что субъект воспринимает этот смешанный сигнал или "реверберируемый сигнал", как хорошее звучание или как
20 естественное звучание. Однако механизм слухового восприятия или механизм, как звук на самом деле воспринимается индивидуумом, сильно нелинеен, не только по отношению к полосам частот, в которых работает человеческий слух, но также и по отношению к обработке сигналов в полосах частот. Кроме того, известно, что человеческое восприятие звука не столько направлено на уровень звукового давления, который
25 может быть вычислен, например, возведением в квадрат цифровых выборок, но восприятие более контролируемо путем ощущения громкости. Кроме того, для смешанных сигналов, которые включают в себя прямую компоненту и реверберационную компоненту сигнала, ощущение громкости реверберационной компоненты зависит не только от вида прямой компоненты сигнала, но также и от
30 уровня или громкости прямой компоненты сигнала.

Таким образом, существует необходимость в определении показателя для воспринимаемого уровня реверберации в сигнале, состоящем из прямой компоненты сигнала и реверберационной компоненты сигнала для того, чтобы справиться с вышеуказанными проблемами, связанными с механизмом слухового восприятия объекта.

35 Задачей настоящего изобретения является, таким образом, обеспечение устройства и способа для определения показателя для воспринимаемого уровня реверберации или обеспечение аудио процессора или способа обработки аудиосигнала с улучшенными характеристиками.

Эта задача достигается устройством для определения показателя воспринимаемого
40 уровня реверберации в соответствии с п.1, способом определения показателя воспринимаемого уровня реверберации в соответствии с п.10, аудио процессором в соответствии с п.11, способом обработки аудиосигнала в соответствии с п.14 или компьютерной программой в соответствии с п.15.

Настоящее изобретение основано на обнаружении того, что показатель для
45 воспринимаемого уровня реверберации сигнала определяется процессором модели громкости, содержащим каскад перцепционного фильтра для фильтрации прямой компоненты сигнала, реверберационной компоненты сигнала или смешанной компоненты сигнала с использованием перцепционного фильтра для моделирования

механизмов слухового восприятия объекта. На основе перцепционно фильтрованных сигналов, блок оценки громкости оценивает первый показатель громкости, используя фильтрованный прямой сигнал и второй показатель громкости, используя

5 фильтрованный реверберационный сигнал или фильтрованный смешанный сигнал. Затем блок объединения объединяет первый показатель и второй показатель, чтобы получить показатель для воспринимаемого уровня реверберации. В частности, путем объединения двух различных показателей громкости, предпочтительно, вычислением разницы, дающим количественное значение или показатель насколько сильно ощущение реверберации по сравнению с ощущением прямого сигнала или смешанного сигнала.

10 Для вычисления показателей громкости могут быть использованы показатели абсолютной громкости и, в частности, показатели абсолютной громкости прямого сигнала, смешанного сигнала или реверберационного сигнала. Кроме того, частичная громкость также может быть вычислена, где первый показатель громкости определяется использованием прямого сигнала в качестве возбуждающего сигнала и

15 реверберационного сигнала в качестве шума в модели громкости, а второй показатель громкости вычисляется использованием реверберационного сигнала в качестве возбуждающего сигнала и прямого сигнала в качестве шума. В частности, объединением этих двух показателей в блоке объединения, получается пригодный показатель для воспринимаемого уровня реверберации. Авторами было обнаружено, что такой

20 пригодный показатель не может быть определен только генерацией одного показателя громкости, например, использованием отдельно прямого сигнала или отдельно смешанного сигнала или отдельно реверберационного сигнала. Вместо этого, из-за взаимозависимости человеческого слуха, объединяя показатели, которые получены отдельно из каждого из этих трех сигналов, воспринимаемый уровень реверберации в

25 сигнале может быть определен или смоделирован с высокой степенью точности. Предпочтительно процессор модели громкости обеспечивает частотно-временное преобразование и подтверждает передаточную функцию уха совместно с моделью возбуждения, фактически происходящей в человеческом слухе, моделируемой слуховыми моделями.

30 В предпочтительном варианте осуществления показатель для воспринимаемого уровня реверберации направляется в блок предсказания, который фактически обеспечивает воспринимаемый уровень реверберации в пригодном масштабе, таком как сон-масштаб. Этот блок предсказания предпочтительно обучен данными теста прослушивания, и параметры блока предсказания для предпочтительно линейного

35 блока предсказания содержат свободный член и коэффициент масштабирования. Свободный член предпочтительно зависит от характеристики фактически используемого фильтра реверберации, и в одном из вариантов осуществления фильтра реверберации характеристический параметр T_{60} , который может быть предоставлен для простых хорошо известных фильтров реверберации, использован в искусственных

40 ревербераторах. Даже если, однако, эта характеристика не известна, например, когда реверберационная компонента сигнала отдельно не доступна, но была отделена от смешанного сигнала до обработки в устройстве по настоящему изобретению, оценка для свободного члена может быть получена.

45 Далее предпочтительные варианты осуществления настоящего изобретения описаны со ссылками на прилагаемые чертежи, на которых:

Фиг.1 является блок-схемой устройства или способа для определения показателя для воспринимаемого уровня реверберации;

Фиг.2а является иллюстрацией предпочтительного варианта осуществления

процессора модели громкости;

Фиг.2b иллюстрирует другой предпочтительный вариант осуществления процессора модели громкости;

5 Фиг.3 иллюстрирует другой предпочтительный вариант осуществления процессора модели громкости;

Фиг.4a,b иллюстрируют примеры огибающих сигнала-времени и соответствующей громкости и частичной громкости;

Фиг.5a,b иллюстрируют информацию экспериментальных данных для обучения блока предсказания;

10 Фиг.6 иллюстрирует блок-схему процессора искусственной реверберации;

Фиг.7 иллюстрирует три таблицы для индикации параметров оценки для вариантов осуществления изобретения;

Фиг.8 иллюстрирует процессор аудиосигнала, реализованный для использования показателя воспринимаемого уровня реверберации с целью искусственной реверберации;

15 Фиг.9 иллюстрирует предпочтительный вариант осуществления блока предсказания, основанный на усредненных по времени воспринимаемых уровнях реверберации; и

Фиг.10 иллюстрирует уравнения из Moore Glasberg, Ваer публикации 1997 года, используемые в предпочтительном варианте осуществления для расчета удельной громкости.

20 Воспринимаемый уровень реверберации зависит как от входного звукового сигнала, так и от импульсной характеристики. Варианты осуществления настоящего изобретения направлены на определение значений этих результатов наблюдения и предсказание воспринимаемого уровня поздней реверберации на основе отдельных сигнальных трактов прямого и реверберирующего сигналов, как они проявляются в цифровых
25 аудио эффектах. Подход к проблеме разработан и впоследствии расширен с учетом влияния времени реверберации на предсказание результата. Это приводит к линейной регрессионной модели с двумя входными переменными, которая способна предсказать воспринимаемый уровень с высокой точностью, как показано на экспериментальных данных, полученных из тестов прослушивания. Вариации этой модели с различными
30 степенью сложности и вычислительной сложностью сравниваются относительно их точности. Применения включают в себя управление цифровыми аудио эффектами для автоматического смешивания аудиосигналов.

Варианты осуществления настоящего изобретения не только пригодны для предсказания воспринимаемого уровня реверберации в речи и музыке, когда прямой
35 сигнал и импульсная характеристика реверберации (ИХР) доступны отдельно. В других вариантах осуществления, в которых возникает реверберируемый сигнал, настоящее изобретение может быть применено также. В этом случае, однако, прямой/пространственный или прямой/реверберационный сепаратор будет включен для отделения прямой компоненты сигнала и реверберационной компоненты сигнала из
40 смешанного сигнала. Такой аудио процессор затем будет полезен для изменения соотношения прямой/реверберационный в этом сигнале для того, чтобы генерировать лучшее звучание реверберационного сигнала или лучшее звучание смешанного сигнала.

Фиг.1 иллюстрирует устройство для определения показателя воспринимаемого уровня реверберации в смешанном сигнале, содержащем прямую компоненту сигнала или «сухую» компоненту 100 сигнала и реверберационную компоненту 102 сигнала.
45 «Сухая» компонента 100 сигнала и реверберационная компонента 102 сигнала вводятся в процессор 104 модели громкости. Процессор модели громкости сконфигурирован для получения прямой компоненты 100 сигнала и реверберационной компоненты 102

сигнала и включает в себя, кроме того, каскад 104a перцепционного фильтра и подключенный затем вычислитель 104b громкости, как проиллюстрировано на фиг.2a. Процессор модели громкости генерирует на своем выходе первый показатель 106 громкости и второй показатель 108 громкости. Оба показателя громкости вводятся в блок 110 объединения для объединения первого показателя 106 громкости и второго показателя 108 громкости чтобы, наконец, получить показатель 112 для воспринимаемого уровня реверберации. В зависимости от реализации, показатель для воспринимаемого уровня 112 может быть введен в блок 114 предсказания для предсказания воспринимаемого уровня реверберации на основе среднего значения по меньшей мере двух показателей для воспринимаемой громкости для разных сигналов, как будет описано в контексте фиг.9. Тем не менее, блок 114 предсказания на фиг.1 необязателен и фактически превращает параметр воспринимаемого уровня в некоторый диапазон значений или диапазон единиц, такой как диапазон сон-единиц, который является пригодным для предоставления количественных значений, связанных с громкостью. Тем не менее, другие коэффициенты для показателя для воспринимаемого уровня 112, которые не обрабатываются блоком 114 предсказания могут также использоваться, например, в аудио процессоре по фиг.8, который не обязательно должен опираться на выходные значения блока 114 предсказания, но который также может непосредственно обрабатывать параметр воспринимаемого уровня 112, либо в прямой форме либо предпочтительно в виде сглаженной формы, где сглаживание с течением времени является предпочтительным для того, чтобы не иметь сильно изменяющиеся коррекции уровня реверберированного сигнала или, как обсуждается позже, коэффициента g усиления, проиллюстрированного на фиг.6 или проиллюстрированного на фиг.8.

В частности, каскад перцепционного фильтра сконфигурирован для фильтрации прямой компоненты сигнала, реверберационной компоненты сигнала или смешанной компоненты сигнала, при этом каскад перцепционного фильтра сконфигурирован для моделирования механизма слухового восприятия объекта, такого как человек, чтобы получить фильтрованный прямой сигнал, фильтрованный реверберационный сигнал или фильтрованный смешанный сигнал. В зависимости от реализации, каскад перцепционного фильтра может включать в себя два фильтра, работающих параллельно или может включать в себя запоминающее устройство и один фильтр, так как один и тот же фильтр действительно может быть использован для фильтрации каждого из трех сигналов, т.е. реверберационного сигнала, смешанного сигнала и прямого сигнала. При этом, однако, следует отметить, что, несмотря на то, что фиг.2a иллюстрирует n фильтров моделирования механизма слухового восприятия, на самом деле двух фильтров будет достаточно, или одного фильтра фильтрующего два сигнала из группы, содержащей реверберационную компоненту сигнала, смешанную компоненту сигнала и прямую компоненту сигнала.

Вычислитель 104b громкости или блок оценки громкости сконфигурирован для оценки первого связанного с громкостью показателя с использованием фильтрованного прямого сигнала и для оценки второго показателя громкости с использованием фильтрованного реверберационного сигнала или фильтрованного смешанного сигнала, где смешанный сигнал получают из суперпозиции прямой компоненты сигнала и реверберационной компоненты сигнала.

Фиг.2c иллюстрирует четыре предпочтительных режима вычисления показателя для воспринимаемого уровня реверберации. Вариант осуществления 1 опирается на частичную громкость, где обе, прямая компонента x сигнала и реверберационная

компонента g сигнала, используются в процессоре модели громкости, но где в целях определения первого показателя EST1 реверберационный сигнал используется в качестве сигнала возбуждения, а прямой сигнал используется в качестве шума. Для определения второго показателя EST2 громкости, ситуация изменилась, и прямая компонента сигнала используется в качестве сигнала возбуждения, а реверберационная компонента сигнала используется в качестве шума. Итак, показатель для воспринимаемого уровня коррекции, сгенерированный блоком объединения, является разностью между первым показателем EST1 громкости и вторым показателем EST2 громкости.

Однако, кроме того имеются другие вычислительно эффективные варианты осуществления, которые показаны в строках 2, 3 и 4 на фиг.2с. Эти более вычислительно эффективные показатели полагаются на вычисление общей громкости трех сигналов, включающих в себя смешанный сигнал m , прямой сигнал x и реверберационный сигнал p . Зависимость от требуемых вычислений, выполняемых блоком объединения, указана в последнем столбце фиг.2с, первым показателем EST1 громкости является общая громкость смешанного сигнала или реверберационного сигнала, а вторым показателем EST2 громкости является общая громкость прямой компоненты x сигнала или смешанной компоненты m сигнала, где фактические объединения проиллюстрированы на фиг.2с.

В дополнительном варианте осуществления процессор модели громкости 104 функционирует в частотной области, как обсуждается более подробно на фиг.3. В такой ситуации, процессор модели громкости и, в частности, вычислитель 104b громкости обеспечивает первый показатель и второй показатель для каждой полосы частот. Эти первые показатели по всем n полосам частот потом суммируют или объединяют вместе в сумматоре 104с для первой ветви и 104d для второй ветви для того, чтобы в конце концов получить первый показатель для широкополосного сигнала и второй показатель для широкополосного сигнала.

Фиг.3 иллюстрирует предпочтительный вариант осуществления процессора модели громкости, который уже обсуждался в некоторых аспектах по отношению к фиг.1, 2a, 2b, 2с. В частности, каскад 104a перцепционного фильтра включает в себя частотно-временной преобразователь 300 для каждой ветви, где, в варианте осуществления по фиг.3, $x[k]$ означает сигнал возбуждения, а $n[k]$ означает шум. Сигнал, преобразованный по времени/частоте, передается в блок 302 передаточной функции уха (Обратите внимание, что передаточная функция уха альтернативно может быть вычислена до частотно-временного преобразователя с аналогичными результатами, но более высокой вычислительной нагрузкой) и выходной сигнал этого блока 302 является входным сигналом в блок 304 вычисления модели возбуждения и идущего следом блока 306 временной интеграции. Затем, в блоке 308 в данном варианте осуществления вычисляется удельная громкость, где блок 308 соответствует блоку 104b вычисления громкости на фиг.2a. Далее, выполняется интегрирование по частоте в блоке 310, где блок 310 соответствует сумматору, уже описанному как 104с и 104d на фиг.2b. Следует отметить, что блок 310 генерирует первый показатель для первого комплекта возбуждающего сигнала и шума, и второй показатель для второго комплекта возбуждающего сигнала и шума. В частности, когда рассматривается фиг.2b, сигналом возбуждения для вычисления первого показателя является реверберационный сигнал, а шумом является прямой сигнал, в то время как для вычисления второго показателя, ситуация изменилась, и сигналом возбуждения является прямая компонента сигнала, а шумом - реверберационная компонента сигнала. Таким образом, для генерации двух различных параметров громкости, процедуры, проиллюстрированные на фиг.3 были выполнены

дважды. Однако изменения в вычислении происходят только в блоке 308, который работает по-разному, как обсуждалось, кроме того, в контексте фиг.10, так что этапы, проиллюстрированные блоками 300 до 306, должны быть выполнены только один раз, и результат временной интеграции блока 306 может быть сохранен для того, чтобы

5 вычислить первую оцененную громкость и вторую оцененную громкость для варианта осуществления 1 на фиг.2с. Следует отметить, что в других вариантах осуществления 2, 3, 4 на фиг.2с, блок 308 заменен отдельным блоком «вычисления общей громкости» для каждой ветви, где, в данном варианте осуществления это безразлично, будь то один сигнал, рассмотренный как возбуждающий сигнал или как шум.

10 Далее более подробно обсуждается модель громкости, проиллюстрированная на фиг.3.

Реализация модели громкости на фиг.3 следует описаниям в [11, 12] с модификациями, детализированными позже. Обучение и проверка предсказания используют данные из тестов прослушивания, описанных в [13] и кратко резюмированных позже. Применение

15 модели громкости для предсказания воспринимаемого уровня поздней реверберации также описано позже. Экспериментальные результаты следуют ниже.

В этом разделе описывается реализация модели частичной громкости, данные теста прослушивания, которые были использованы в качестве экспериментальных данных для вычислительного предсказания воспринимаемого уровня реверберации, и

20 предлагаемый способ предсказания, который основан на модели частичной громкости.

Модель громкости вычисляет частичную громкость $N_{x,n}[k]$ сигнала $x[k]$, который представлен вместе с маскирующим сигналом $n[k]$

$$N_{x,n}[k] = f(x[k], n[k]). \quad (1)$$

25 Хотя ранние модели имели дело с восприятием громкости в установившемся фоновом шуме, существуют некоторые работы по восприятию громкости в фонах совместно модулированных случайных шумов [14], комплексных внешних звуков [12], и музыкальных сигналов [15]. Фиг.4b иллюстрирует общую громкость и частичную громкость компонент примерного сигнала, показанного на фиг.4а, вычисленных с

30 моделью громкости, используемой здесь.

Модель, используемая в этой работе, аналогична модели в [11, 12], которая сама выведена в более раннем исследовании Fletcher, Munson, Stevens, и Zwicker, с некоторыми изменениями, как описано ниже. Блок-схема модели громкости показана на фиг.3. Входные сигналы обрабатываются в частотной области с использованием

35 кратковременного преобразования Фурье (КВПФ). В [12], 6 ДПФ (дискретное преобразование Фурье) различной длины используются для того, чтобы получить хорошее совпадение для разрешения по частоте и временного разрешения с разрешением по частоте и временным разрешением человеческой слуховой системы на всех частотах. В этой работе используется только ДПФ одной длины ради вычислительной

40 эффективности, с длиной кадра 21 мс при частоте дискретизации 48 кГц, 50% перекрытием и функцией окна Ханна. Передача через наружное и среднее ухо моделируется с фиксированным фильтром. Функция возбуждения вычисляется для 40 полос частот слухового фильтра расположенных на шкале эквивалентного

45 прямоугольного диапазона частот (ЭПДЧ) с использованием модели зависимости от уровня возбуждения. В дополнение к временной интеграции из-за обработки методом окна КВПФ, рекурсивная интеграция реализуется с постоянной времени 25 мс, которая активна только в то время, когда затухает сигнал возбуждения.

Удельная частичная громкость, т.е. частичная громкость, вызванная в каждой полосе частот слухового фильтра, вычисляется из уровней возбуждения от важного сигнала

(возбуждающего сигнала) и звуковых помех в соответствии с уравнениями (17)-(20) в [11], проиллюстрированных на фиг.10. Эти уравнения охватывают четыре случая, где сигнал выше порога слышимости в шуме или нет, и где возбуждение смешанного сигнала меньше 100 дБ или нет. Если никакой сигнал помехи не подается в модель, т.е. $n[k]=0$, результат равен общей громкости $N_x[k]$ сигнала возбуждения $x[k]$.

В частности, фиг.10 иллюстрирует уравнения 17, 18, 19, 20 публикации "A Model for the Prediction of Thresholds, Loudness and Partial Loudness", В.С.Дж. Мур, В.Р. Глазберг, Т. Ваер, J. Audio Eng. Soc, изд. 45, No. 4, апрель 1997. Эта ссылка описывает случай сигнала, представленного вместе с фоновым звуком. Несмотря на то, что фон может быть любым типом звука, он упоминается как "шум" в этой ссылке, чтобы отличить его от сигнала, громкость которого оценивается. Наличие шума уменьшает громкость сигнала, эффект называется частичной маскировкой. Громкость сигнала растет очень быстро, когда его уровень увеличивается от порогового значения до значения 20-30 дБ выше порога. В публикации предполагается, что частичная громкость сигнала, представленная в шуме, может быть вычислена путем суммирования удельных частичных громкостей сигнала по частоте (на ЭПДЧ-шкале). Уравнения, полученные для вычисления удельной частичной громкости, рассматривают четыре предельных случаях. E_{SIG} обозначает возбуждение вызванное сигналом, а E_{NOISE} обозначает возбуждение, вызванное шумом.

Предполагается, что $E_{SIG} > E_{THRQ}$ и E_{SIG} плюс $E_{NOISE} < 10^{10}$. Общая удельная громкость N'_{TOT} определяется следующим образом:

$$N'_{TOT} = C\{[(E_{SIG} + E_{NOISE})G + A]^a - A^a\}$$

Предполагается, что слушатель может разделить удельную громкость при заданной центральной частоте между удельной громкостью сигнала и удельной громкостью шума, но таким образом, что предпочитает общую удельную громкость.

$$N'_{TOT} = N'_{SIG} + N_{NOISE}$$

Это предположение согласуется, так как в большинстве экспериментов по измерению частичного маскирования, слушатель слышит первым один только шум и затем шум плюс сигнал. Удельная громкость только для шума при условии, что он находится выше порога, будет

$$N'_{NOISE} = C\{[(E_{NOISE})G + A]^a - A^a\}.$$

Таким образом, если удельная громкость сигнала была получена только зависимостью удельной громкости шума от общей удельной громкости, результатом будет

$$N'_{SIG} = C\{[(E_{SIG} + E_{NOISE})G + A]^a - A^a\} - C\{[(E_{NOISE})G + A]^a - A^a\}$$

На практике, способ, удельная громкость которого распределена между сигналом и шумом, по-видимому меняется в зависимости от относительного возбуждения сигнала и шума.

Рассмотренные четыре ситуации показывают как удельная громкость определяется на разных уровнях сигнала. Пусть E_{THRQ} обозначает пиковое возбуждение, вызванное синусоидальным сигналом, когда он является маскирующим порогом в фоновом шуме. Когда E_{SIG} значительно ниже E_{THRQ} , вся удельная громкость присвоена шуму, а частичная удельная громкость сигнала приближается к нулю. Во-вторых, когда E_{NOISE} значительно ниже E_{THRQ} , частичная удельная громкость приближается к значению, которое она будет иметь для сигнала в тишине. В-третьих, когда сигнал находится на его маскирующем пороге, при возбуждении E_{THRQ} , предполагается, что частичная

удельная громкость будет равна значению, которое возникнет для сигнала на абсолютном пороге. Наконец, когда сигнал находится в центре узкополосного шума значительно выше его маскирующего порога, громкость сигнала приближается к своему демаскирующему значению. Поэтому частичная удельная громкость сигнала также приближается к своему демаскирующему значению.

Рассмотрим последствия этих различных граничных условий. На маскирующем пороге, удельная громкость равна той, что для сигнала на пороге в тишине. Эта удельная громкость меньше, чем можно было бы предсказать из приведенного выше уравнения, по-видимому, потому что некоторая удельная громкость сигнала приписывается к шуму. Для того, чтобы получить правильную удельную громкость для сигнала, предполагается, что удельная громкость, приписанная к шуму, увеличивается фактором B , где

$$B = \frac{[(E_{\text{THR}} + E_{\text{NOISE}})G + A]^a - (E_{\text{THR}}G + A)^a}{E_{\text{NOISE}}G + A)^a - A^a}$$

Применение этого фактора ко второму члену в вышеизложенном уравнении для N'_{SIG} дает

$$N'_{\text{SIG}} = C\{[(E_{\text{SIG}} + E_{\text{NOISE}})G + A]^a - A^a\} - C\{[(E_{\text{THR}} + E_{\text{NOISE}})G + A]^a - (E_{\text{THR}}G + A)^a\}$$

Предполагается, что, когда сигнал на маскирующем пороге, его пиковое возбуждение E_{THR} равно $KE_{\text{NOISE}} + E_{\text{THRQ}}$, где K отношение сигнал-шум на выходе слухового фильтра, требующегося для порога на более высоких уровнях маскирования. Недавние оценки K , полученные для маскировки экспериментов, использующих шум с узкополосным провалом в спектре, предполагают, что K заметно возрастает на очень низких частотах, становится больше единицы. В этой ссылке значение K оценивается как функция частоты. Значение уменьшается от высоких уровней на низких частотах к неизменным низким уровням на более высоких частотах. К сожалению, нет оценки для K для центральных частот ниже 100 Гц, поэтому значения от 50 до 100 Гц заменяют E_{THR} в вышеприведенных результатах уравнения в:

$$N'_{\text{SIG}} = C\{[(E_{\text{SIG}} + E_{\text{NOISE}})G + A]^a - A^a\} - C\{[(E_{\text{NOISE}}(1 + K) + E_{\text{THRQ}})G + A]^a - (E_{\text{THRQ}}G + A)^a\}$$

Когда $E_{\text{SIG}} = E_{\text{THR}}$ это уравнение устанавливает пиковую удельную громкость для сигнала на абсолютном пороге в тишине.

Когда сигнал намного выше его маскирующего порога, то есть, когда $E_{\text{SIG}} \gg E_{\text{THR}}$, удельная громкость сигнала приближается к значению, которое она будет иметь, когда не присутствуют фоновые шумы. Это означает, что удельная громкость, приписанная к шуму, становится исчезающе малой. Чтобы обеспечить это, вышеприведенное уравнение модифицировано путем введения дополнительного члена, который зависит от отношения $E_{\text{THR}}/E_{\text{SIG}}$. Этот член уменьшается, как только E_{SIG} повысится выше значения, соответствующего маскирующему порогу. Таким образом, вышеприведенное уравнение становится уравнением 17 на фиг.10.

Это окончательное уравнение для N'_{SIG} в случае, когда $E_{\text{SIG}} > E_{\text{THR}}$ и $E_{\text{SIG}} + E_{\text{NOISE}} \leq 10^{10}$. Показатель 0,3 в конечном члене был выбран эмпирически так, чтобы получить хорошее соответствие данных громкости тона в шуме в зависимости от отношения сигнал-шум.

Потом рассматривается ситуация, где $E_{\text{SIG}} < E_{\text{THR}}$. В предельном случае, когда E_{SIG} чуть ниже E_{THR} , удельная громкость приблизилась бы к значению, указанному в уравнении 17 на фиг.10. Когда E_{SIG} снижается до значения значительно ниже E_{THR} ,

удельная громкость должна быстро стать очень малой. Это достигается уравнением 18 на фиг.10. Первый член в скобках определяет скорость, с которой уменьшается удельная громкость E_{SIG} , снижается ниже E_{THRN} . Это описывает зависимость между удельной громкостью и возбуждением для сигнала в тишине, когда $E_{SIG} < E_{THRN}$,
 5 исключая что E_{THRN} была заменена в уравнении 18. Первый член в фигурных скобках обеспечивает приближение удельной громкости к значению, определяемому уравнением 17 фиг.10, как E_{SIG} приближающуюся к E_{THRN} .

Описанные уравнения для частичной громкости до сих пор применяются, когда
 10 $E_{SIG} + E_{NOISE} < 10^{10}$. Применением тех же рассуждений, которые использовались для вывода уравнения (17) на фиг.10, может быть получено любое уравнение для случая $E_{NOISE} \geq E_{THRN}$ и $E_{SIG} + E_{NOISE} > 10^{10}$, как указано в уравнении 19 на фиг.10. $C_2 = C / (1,04 \cdot 10^6)^{0,5}$. Аналогично, применением тех же рассуждений, используемых для вывода
 15 уравнения (18) на фиг.30, может быть получено уравнение для случая, когда $E_{SIG} < E_{THRN}$ и $E_{SIG} + E_{NOISE} > 10^{10}$, как указано в уравнении 20 на фиг.10.

Следующие пункты следует отметить. Эта известная модель применяется для
 настоящего изобретения, где на первом этапе, SIG соответствует, например, прямому
 20 сигналу в качестве "возбуждающего сигнала" и шуму соответствует, например, реверберационный сигнал или смешанный сигнал в качестве "шума". На втором этапе, как обсуждалось в контексте первого варианта осуществления на фиг.2с, SIG будет тогда соответствовать реверберационному сигналу в качестве "возбуждающего сигнала" и "шум" будет соответствовать прямому сигналу. Затем получают два показателя
 25 громкости, которые затем объединяются в блоке объединения предпочтительно формированием разности.

Для того, чтобы оценить пригодность описанной модели громкости для задач
 предсказания воспринимаемого уровня поздней реверберации, массив
 экспериментальных данных, сгенерированный из ответов слушателя является
 30 предпочтительным. С этой целью данные исследования, характеризующие несколько тестов прослушивания [13], используются в данном документе, которые кратко сформулированы ниже. Каждый тест прослушивания состоял из множественных экранов графического пользовательского интерфейса, которые представлены смещениями
 35 разных прямых сигналов с разными условиями искусственной реверберации. Слушателям было предложено оценить это воспринимаемую величину реверберации по шкале от 0 до 100 баллов. Кроме того, два сигнала привязки были представлены на 10 баллах и на 90 баллах. Сигналы привязки были созданы из того же прямого сигнала с разными условиями реверберации.

Прямыми сигналами, используемыми для создания элементов теста, были
 40 монофонические записи речи, отдельных инструментов и музыки различных жанров с длиной около 4 секунд каждый. Было использовано большинство из элементов, происходящих от безэховой записи, а также коммерческие записи с малой величиной оригинальной реверберации.

ИХР представляют позднюю реверберацию и были сгенерированы с использованием
 45 экспоненциально затухающего белого шума с частотно-зависимыми скоростями затухания. Скорости затухания выбраны так, что время реверберации уменьшается от низких до высоких частот, начиная с основного времени реверберации T_{60} . Ранними отражениями пренебрегли в этой работе. Реверберационный сигнал $r[k]$ и прямой сигнал

$x[k]$ были измерены и суммированы так, что отношение их средних показателей громкости в соответствии с ITU-R BS.1770 [16] соответствует желаемой ОПР и так, что все смещения тестового сигнала имеют равную долгосрочную громкость. Все участники испытаний работали в области аудио и имели опыт работы с субъективными тестами прослушивания.

Экспериментальные данные, используемые для обучения и проверки/тестирования способа предсказания были взяты из двух тестов прослушивания и обозначены А и В, соответственно. Набор данных А состоял из оценок 14 слушателей для 54 сигналов. Слушатели повторили тест один раз и средняя оценка была получена от всех из 28 оценок для каждого элемента. 54 сигнала были сгенерированы объединением 6 различных прямых сигналов и 9 условий стереофонической реверберации, с $T_{60} \in \{1, 1,6, 2,4\}$ с и ОПР $\in \{3, 7,5, 12\}$ дБ, и без предварительной задержки.

Данные в В были получены из оценок 14 слушателей для 60 сигналов. Сигналы были получены использованием 15 прямых сигналов и 36 условий реверберации. Условия реверберации выбраны по четырем параметрам, а именно T_{60} , ОПР, предварительная задержка, и МКК. Для каждого прямого сигнала 4 ИХР были выбраны так, что два не имеют предварительной задержки и два имеют короткую предварительную задержку в 50 мс, и два были монофоническими, а два были стереофоническими.

Далее обсуждаются дополнительные характеристики предпочтительного варианта осуществления блока 110 объединения на фиг.1.

Основной входной характеристикой для способа предсказания является вычисленная разность частичной громкости $N_{r,x}[k]$ реверберационного сигнала $r[k]$ (с прямым сигналом $x[k]$ являющимся помехой) и громкости $N_{x,r}[k]$ сигнала $x[k]$ (где $r[k]$ является помехой) в соответствии с уравнением 2.

$$\Delta N_{r,x}[k] = N_{r,x}[k] - N_{x,r}[k] \quad (2)$$

Смысл уравнения (2) в том, что разность $\Delta N_{r,x}[k]$ является показателем того, насколько сильно ощущение реверберации по сравнению с ощущением прямого сигнала. Взятие разности было также установлено, чтобы сделать результат предсказания примерно инвариантным по отношению к уровню воспроизведения. Уровень воспроизведения оказывает воздействие на исследуемое ощущение [17, 8], но в более тонкой степени, чем свидетельствует увеличение частичной громкости $N_{r,x}$ с увеличением уровня воспроизведения. Как правило, музыкальные записи звучат более реверберирующими на средних и высоких уровнях (начиная примерно от 75-80 дБ), чем при примерно от 12 до 20 дБ нижних уровней. Этот эффект особенно очевиден в тех случаях, когда ОПР является положительным, что действительно «почти для всей записанной музыки» [18], но не во всех случаях для концертной музыки, где «слушатели находятся часто далеко за пределами критического расстояния» [6].

Уменьшение воспринимаемого уровня реверберации с уменьшением уровня воспроизведения лучше всего объясняется тем, что динамический диапазон реверберации меньше, чем прямые звуки (или частотно-временное представление реверберации более плотное, тогда как частотно-временное представление прямых звуков более редкое [19]). В таком сценарии реверберационный сигнал, скорее всего, ниже порога слышимости, чем прямые звуки.

Несмотря на то, что уравнение (2) описывает, в качестве операции объединения, разность между двумя показателями громкости $N_{r,x}[k]$ и $N_{x,r}[k]$, могут быть также выполнены другие объединения, такие как умножения, деления или даже дополнения.

В любом случае, достаточно, чтобы две альтернативы, указанные двумя показателями громкости, были объединены для того, чтобы в результате иметь влияние обоих альтернатив. Однако эксперименты показали различные результаты в наилучших значениях из модели, т.е. в результатах модели, которые вписываются в тесты прослушивания в хорошей степени, разность является предпочтительным способом объединения.

Далее описаны подробности блока 114 предсказания, иллюстрированного на фиг.1, где эти подробности относятся к предпочтительным вариантам осуществления.

Способы предсказания, описанные ниже, являются линейными и используют подбор методом наименьших квадратов для вычисления коэффициентов модели. Простая структура предсказания является предпочтительной в случаях, когда размер наборов данных для обучения и тестирования предсказания ограничен, что может привести к подгонке модели при использовании методов регрессии с большими степенями свободы, например, нейронных сетей. Базовое предсказание R_b получается путем линейной регрессии в соответствии с уравнением (3) с коэффициентами a_1 , с K , являющимся длиной сигнала в кадрах,

$$\hat{R}_b = a_0 + a_1 \frac{1}{K} \sum_{k=1}^K \Delta N_{r,x}[k]. \quad (3)$$

Модель имеет только одну независимую переменную, т.е. среднее $\Delta N_{r,x}[k]$. Для отслеживания изменений и чтобы иметь возможность выполнять обработку в реальном времени, вычисление среднего может быть аппроксимировано с использованием квазиинтегратора. Модель параметров, полученных с использованием комплекта данных A для обучения $a_0=48,2$ и $a_1=14,0$, где a_0 равно средней оценке для всех слушателей и элементов.

Фиг.5а изображает предсказания ощущений для комплекта данных A . Видно, что предсказания умеренно коррелируют со средними оценками слушателя с коэффициентом корреляции 0,71. Пожалуйста, обратите внимание, что выбор коэффициентов регрессии не влияет на эту корреляцию. Как показано на нижнем графике для каждого смещения, генерируемого теми же прямыми сигналами, точки показывают характерную форму по центру близко к диагонали. Эта форма указывает, что, хотя базовая модель \hat{R}_b может предсказать R в некоторой степени, она не отражает влияние T_{60} на оценки. Визуальный осмотр точек данных предполагает линейную зависимость от T_{60} . Если значение T_{60} известно, как в случае при управлении звуковым эффектом, оно может быть легко включено в модель линейной регрессии для получения улучшенного предсказания

$$\hat{R}_e = a_0 + a_1 \frac{1}{K} \sum_{k=1}^K \Delta N_{r,x}[k] + a_2 T_{60}. \quad (4)$$

Параметры модели, полученные из комплекта данных A являются $a_0=48,2$, $a_1=14,0$, $a_2=10,2$. Результаты показаны на фиг.5b отдельно для каждого комплекта данных.

Оценка результатов будет описана более подробно в следующем разделе.

Кроме того, усреднение по блокам более или менее может быть выполнено, пока происходит усреднение, по меньшей мере, по двум блокам, хотя, в связи с теорией линейных уравнений, лучшие результаты могут быть получены, когда выполняется усреднение по всей музыкальной части до некоторого кадра. Однако для применений в реальном времени предпочтительно сократить количество кадров, по которым усредняют, в зависимости от конкретного применения.

Фиг.9 дополнительно иллюстрирует, что свободный член задан a_0 и $a_2 \cdot T_{60}$. Второй

член $a_2 \cdot T_{60}$ был выбран для того, чтобы быть в состоянии применить это уравнение не только к одному ревербератору, т.е. к ситуации, в которой фильтр 600 по фиг.6 не изменяется. Это уравнение, которое, конечно, является свободным членом, но которое зависит от фактически использованных фильтров 606 реверберации фиг.6, обеспечивает, таким образом, возможность использовать точно такое же уравнение для других фильтров реверберации, имеющих другие значения T_{60} . Как известно из уровня техники, T_{60} это параметр, описывающий определенный фильтр реверберации и, в частности, означает, что энергия реверберации была уменьшена на 60 дБ от начального максимального значения энергии реверберации. Как правило, кривые реверберации уменьшаются со временем, и поэтому T_{60} указывает период времени, в котором энергия реверберации, генерируемая сигналом возбуждения, уменьшилась на 60 дБ. Аналогичные результаты с точки зрения точности предсказания получаются заменой T_{60} на параметры, представляющие аналогичную информацию (т.е. длину ИХР), например, T_{30} .

Ниже модели оцениваются с использованием коэффициента r корреляции, средней абсолютной ошибки (САО) и среднеквадратической ошибки (СКО) между средними оценками слушателя и предсказанным ощущением. Эксперименты выполнены в виде двойной перекрестной проверки, то есть блок предсказания обучен комплектом данных А и протестирован комплектом данных В, и эксперимент повторялся с В для обучения и А для тестирования. Оценки показателей, полученные от обоих прохождений, усредняются отдельно для обучения и тестирования.

Результаты показаны в таблице 1 для моделей предсказания \hat{R}_s и \hat{R}_e . Параметр \hat{R}_e предсказания дает точные результаты с СКО на 10,6 балла. Среднее стандартное отклонение индивидуальных оценок слушателя на элемент приведены в качестве показателя дисперсии от среднего (от оценок всех слушателей на элемент) как $\overline{\sigma}_A = 13,4$ для комплекта данных А и $\overline{\sigma}_B = 13,6$ для комплекта данных В. Сравнение с СКО указывает, что \hat{R}_e , по крайней мере, точно, как средний слушатель в тесте прослушивания.

Точности предсказаний для комплектов данных немного различаются, например, для \hat{R}_e обоих САО и СКО примерно на один балл ниже среднего значения (как указано в таблице) при тестировании с комплектом данных А и на один балл выше среднего при тестировании с комплектом данных В. Тот факт, что оценки показателей для обучения и тестирования сопоставимы, указывает на то, что переобучения блока предсказания удалось избежать.

Для того, чтобы облегчить экономическую реализацию таких моделей предсказания, следующие эксперименты исследуют как использование характеристик громкости с меньшей вычислительной сложностью влияет на точность предсказания результата. Эксперименты направлены на замену вычисления частичной громкости оценками общей громкости и на упрощенную реализацию модели возбуждения.

Вместо использования разности частичной громкости $\Delta N_{r,x}[k]$ исследованы оценки трех разностей общей громкости, с громкостью прямого сигнала $N_x[k]$, громкостью реверберации $N_r[k]$, и громкостью смешанного сигнала $N_m[k]$, как показано в уравнениях (5)-(7) соответственно.

$$\Delta N_{m-x}[k] = N_m[k] - N_x[k] \quad (5)$$

Уравнение (5) основано на предположении, что воспринимаемый уровень реверберационного сигнала может быть выражен как разность (увеличение) общей

громкости, которая вызвана добавлением реверберации в «сухой» сигнал.

Ниже следует похожее обоснование как для разности частичной громкости в уравнении (2), Характеристики громкости использованием разностей общей громкости реверберационного сигнала и смешанного сигнала или прямого сигнала, соответственно, определено в уравнениях (6) и (7). Показатель для предсказания ощущения происходит от громкости реверберационного сигнала при прослушивании отдельно, с вычитанием членов для моделирования частичного маскирования и для нормализации по отношению уровня воспроизведения, полученного из смешанного сигнала или прямого сигнала, соответственно.

$$\Delta N_{r-m}[k] = N_r[k] - N_m[k] \quad (6)$$

$$\Delta N_{r-x}[k] = N_r[k] - N_x[k] \quad (7)$$

В таблице 2 показаны результаты, полученные с характеристиками, основанными на общей громкости и показывающими, что фактически две из них, $\Delta N_{m-x}[k]$ и $\Delta N_{r-x}[k]$, выдают предсказания почти с той же точностью, как \hat{R}_e . Однако, как показано в таблице 2, даже $\Delta N_{r-n}[k]$ предусматривает использование для результатов.

Наконец, в дополнительном эксперименте, исследовано влияние реализации функции рассеяния. Это имеет особое значение для многих сценариев применения, потому что использование уровня зависимости моделей возбуждения требует реализаций высокой вычислительной сложности. Эксперименты с аналогичной обработкой, как для \hat{R}_e , но с использованием одной модели громкости без рассеяния и одной модели громкости с инвариантной по уровням функцией рассеяния, привели к результатам, показанным в таблице 2. Влияние рассеяния представляется незначительным.

Таким образом, уравнения (5), (6) и (7), которые показывают варианты осуществления 2, 3, 4 по фиг.2с иллюстрируют, что даже без частичных громкостей, но с общими громкостями, для различных объединений компонент сигнала или сигналов, хорошие значения или показатели для воспринимаемого уровня реверберации в смешанном сигнале также получены.

Далее, предпочтительное применение изобретательского определения показателей для воспринимаемого уровня реверберации обсуждается в контексте фиг.8. Фиг.8 иллюстрирует аудио процессор для генерации реверберационного сигнала из прямой компоненты сигнала, введенной на входе 800. Прямая или «сухая» компонента сигнала вводится в ревербератор 801, который может быть похожим на ревербератор 606 на фиг.6. «Сухая» компонента сигнала входа 800 дополнительно вводится в устройство 802 для определения показателя для воспринимаемой громкости, который может быть реализована, как обсуждалось в контексте фиг.1, фиг.2а и 2с, 3, 9 и 10. Выходным сигналом устройства 802 является показатель R для воспринимаемого уровня реверберации в смешанном сигнале, который вводится в контроллер 803. Контроллер 803 получает, еще на входе, целевое значение для показателя воспринимаемого уровня реверберации и вычисляет из этого целевого значения и фактического значения R снова значение на выходе 804.

Это значение усиления вводится в манипулятор 805, который выполнен для обработки, в этом варианте осуществления, реверберационной компоненты 806 сигнала, выведенной ревербератором 801. Как показано на фиг.8, устройство 802 дополнительно принимает реверберационную компоненту 806 сигнала, обсуждавшуюся в контексте с фиг.1 и других фиг, описывающих устройство для определения показателя воспринимаемой громкости. Выходной сигнал манипулятора 805 вводится в сумматор 807, где выходной сигнал манипулятора содержит в варианте осуществления фиг.8

обработанную реверберационную компоненту, а выходной сигнал сумматора 807 указывает смешанный сигнал 808 с воспринимаемой реверберацией, как определено целевым значением. Контроллер 803 может быть выполнен для реализации любого из правил управления, определенных в уровне техники, для управления обратной связью, где целевым значением является заданное значение и значение R, генерируемое устройством, является фактическим значением и усиление 804 выбирается так, чтобы фактическое значение R приближалось к целевому значению, введенному в контроллер 803. Хотя фиг.8 иллюстрирует, что реверберационный сигнал обработан усилением в манипуляторе 805 который, в частности, содержит умножитель или блок весовой обработки, другие реализации могут быть выполнены также. Еще одна реализация, например, состоит в том, что нет реверберационного сигнала 806, но «сухая» компонента сигнала обработана манипулятором, как указано дополнительной линией 809. В этом случае, необработанная реверберационная компонента сигнала в качестве выходного сигнала ревербератора 801 будет введена в сумматор 807, как указано дополнительной линией 810. Естественно, что даже обработка «сухой» компоненты сигнала и реверберационной компоненты сигнала может быть выполнена для того, чтобы ввести или установить некоторую степень воспринимаемой громкости реверберации в смешанном сигнале 808, выведенном сумматором 807. Еще одна реализация, например, состоит в том, что время реверберации T_{60} обрабатывают.

Настоящее изобретение обеспечивает простое и надежное предсказание воспринимаемого уровня реверберации и, в частности, поздней реверберации в речи и музыке, с использованием моделей громкости различной сложности вычислений. Модули предсказания были обучены и оценены с использованием субъективных данных, полученных из трех тестов прослушивания. В качестве отправной точки, использование модели частичной громкости привело к модели предсказания с высокой точностью, когда известен T_{60} из ИХР 606 по фиг.6. Этот результат также интересен с перцепционной точки зрения, когда считается, что модель частичной громкости не была первоначально разработана с возбуждением прямого и реверберационного звука, как обсуждено в контексте фиг.10. Последующие изменения вычисления входных характеристик для способа предсказания приводят к серии упрощенных моделей, которые были показаны для достижения сопоставимой производительности для имеющихся комплектов данных. Эти изменения включали использование моделей общей громкости и упрощенной функций рассеяния. Варианты осуществления настоящего изобретения также применимы для более разнообразных ИХР включая ранние отражения и большие предварительные задержки. Настоящее изобретение также пригодно для определения и управления вкладом воспринимаемой громкости других типов дополнения или реверберационных аудио эффектов.

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

В зависимости от некоторых требований реализации, варианты осуществления изобретения могут быть реализованы в аппаратных средствах или в программном обеспечении. Реализация может быть выполнена с использованием цифрового носителя информации, например, гибкого диска, DVD, CD, ROM, PROM, EPROM, EEPROM или флэш-памяти, имеющих электронным образом считываемые сигналы управления сохраненные на нем, которые взаимодействуют (или способны к взаимодействию) с

программируемой компьютерной системой так, что выполняется соответствующий способ.

Некоторые варианты осуществления настоящего изобретения включают в себя невременные или материальные носители данных имеющие электронным образом считываемые сигналы управления, которые способны взаимодействовать с программируемой компьютерной системой так, что выполняется один из способов, описанных здесь.

Как правило, варианты осуществления настоящего изобретения могут быть реализованы в качестве компьютерного программного продукта с программным кодом, причем программный код функционирует для выполнения одного из способов, когда компьютерный программный продукт выполняется на компьютере. Программный код может, например, быть сохранен на машиночитаемом носителе.

Другие варианты осуществления включают в себя компьютерную программу для выполнения одного из способов, описанных здесь, хранящуюся на машиночитаемом носителе.

Другими словами, вариант осуществления способа согласно изобретению является, таким образом, компьютерной программой, имеющей программный код для выполнения одного из способов, описанных здесь, когда компьютерная программа выполняется на компьютере.

Еще одним вариантом осуществления способов изобретения является, таким образом, носитель данных (или цифровой носитель, или компьютерно-читаемый носитель), содержащий, записанную на нем, компьютерную программу для выполнения одного из способов, описанных здесь.

Еще одним вариантом осуществления способа изобретения является, таким образом, поток данных или последовательность сигналов, представляющих компьютерную программу для выполнения одного из способов, описанных здесь. Поток данных или последовательность сигналов могут, например, быть сконфигурированы для передачи через интерфейс передачи данных, например, через Интернет.

Еще один вариант осуществления включает в себя средства обработки, например, компьютер или программируемое логическое устройство, сконфигурированное или адаптированное для выполнения одного из способов, описанных здесь.

Другой вариант осуществления содержит компьютер, имеющий установленную на него компьютерную программу для выполнения одного из способов, описанных здесь.

В некоторых вариантах осуществления, программируемое логическое устройство (например, программируемая пользователем вентильная матрица) может использоваться для выполнения некоторых или всех функциональных возможностей способов, описанных здесь. В некоторых вариантах осуществления, программируемая пользователем вентильная матрица может взаимодействовать с микропроцессором для выполнения одного из способов, описанных здесь. Вообще, способы предпочтительно выполнены любыми аппаратными средствами устройств.

Описанные выше варианты осуществления являются только иллюстративными для принципов настоящего изобретения. Понятно, что модификации и варианты механизмов и деталей, описанных здесь, будут очевидны для специалистов в данной области техники. Замысел, таким образом, должен быть ограничен только объемом последующей формулы изобретения, а не конкретными деталями, представленными посредством описания и объяснения вариантов осуществления в данном документе.

СПИСОК ЛИТЕРАТУРЫ:

[1] A. Czyzewski, "A method for artificial reverberation quality testing," *J. Audio Eng. Soc.*, vol. 38, pp. 129-141, 1990.

5 [2] J.A. Moorer, "About this reverberation business," *Computer Music Journal*, vol. 3, 1979.

[3] B. Scharf, "Fundamentals of auditory masking," *Audiology*, vol. 10, pp. 30-40, 1971.

[4] W.G. Gardner and D. Griesinger, "Reverberation level matching experiments," in *Proc. of the Sabine Centennial Symposium, Acoust. Soc. of Am.*, 1994.

10 [5] D. Griesinger, "How loud is my reverberation," in *Proc. Of the AES 98th Conv.*, 1995.

[6] D. Griesinger, "Further investigation into the loudness of running reverberation," in *15 Proc. of the Institute of Acoustics (UK) Conference*, 1995.

[7] D. Lee and D. Cabrera, "Effect of listening level and background noise on the subjective decay rate of room impulse responses: Using time varying-loudness to model reverberance," *Applied Acoustics*, vol. 71, pp. 801-811, 2010.

20 [8] D. Lee, D. Cabrera, and W.L. Martens, "Equal reverberance matching of music," *Proc. of Acoustics*, 2009.

[9] D. Lee, D. Cabrera, and W.L. Martens, "Equal reverberance matching of running musical stimuli having various reverberation times and SPLs," in *Proc. of the 20th International Congress on Acoustics*, 2010.

25 [10] A. Tsilfidis and J. Mourjopoulos, "Blind single-channel suppression of late reverberation based on perceptual reverberation modeling," *J. Acoust. Soc. Am*, vol. 129, pp. 1439-1451, 2011.

[11] B.C.J. Moore, B.R. Glasberg, and T. Baer, "A model for the prediction of threshold, loudness, and partial loudness," *J. Audio Eng. Soc.*, vol. 45, pp. 224-240, 1997.

35 [12] B.R. Glasberg and B.C.J. Moore, "Development and evaluation of a model for predicting the audibility of time varying sounds in the presence of the background sounds," *J. Audio Eng. Soc.*, vol. 53, pp. 906-918, 2005.

40 [13] J. Paulus, C. Uhle, and J. Herre, "Perceived level of late reverberation in speech and music," in *Proc. of the AES 130th Conv.*, 2011.

[14] J.L. Verhey and S.J. Heise, "Einfluss der Zeitstruktur des Hintergrundes auf die Tonhaltigkeit und Lautheit des tonalen Vordergrundes (in German)," in *Proc. of DAGA*, 45 2010.

[15] C. Bradter and K. Hobohm, "Loudness calculation for individual acoustical objects within complex temporally variable sounds," in *Proc. of the AES 124th Conv.*, 2008.

5 [16] International Telecommunication Union, Radiocommunication Assembly, "Algorithms to measure audio programme loudness and true-peak audio level," Recommendation ITU-R BS. 1770, 2006, Geneva, Switzerland.

10 [17] S. Hase, A. Takatsu, S. Sato, H. Sakai, and Y. Ando, "Reverberance of an existing hall in relation to both subsequent reverberation time and SPL," *J. Sound Vib.*, vol. 232, pp. 149-155, 2000.

[18] D. Griesinger, "The importance of the direct to reverberant ratio in the perception of distance, localization, clarity, and envelopment," in *Proc. of the AES 126th Conv.*, 2009.

15 [19] C. Uhle, A. Walther, O. Hellmuth, and J. Herre, "Ambience separation from mono recordings using Non-negative Matrix Factorization," in *Proc. of the AES 30th Conf.*, 2007.

Формула изобретения

20 1. Устройство для определения показателя воспринимаемого уровня реверберации в смешанном сигнале, состоящем из прямой компоненты (100) сигнала и реверберационной компоненты (102) сигнала, содержащее:

процессор (104) модели громкости, содержащий каскад перцепционного фильтра для фильтрации «сухой» компоненты (100) сигнала, реверберационной компоненты (102) сигнала или смешанного сигнала, причем каскад перцепционного фильтра выполнен с возможностью моделирования механизма слухового восприятия объекта для получения фильтрованного прямого сигнала, фильтрованного реверберационного сигнала или фильтрованного смешанного сигнала;

30 блок оценки громкости для оценки первого показателя громкости с использованием фильтрованного прямого сигнала и для оценки второго показателя громкости с использованием фильтрованного реверберационного сигнала или фильтрованного смешанного сигнала, где фильтрованный смешанный сигнал получен из суперпозиции прямой компоненты сигнала и реверберационной компоненты сигнала; и

35 блок (110) объединения для объединения первого и второго показателей (106, 108) громкости, чтобы получить показатель (112) для воспринимаемого уровня реверберации.

2. Устройство по п.1, в котором блок (104b) оценки громкости выполнен с возможностью оценки первого показателя громкости так, что фильтрованный прямой сигнал считается возбуждающим сигналом, а фильтрованный реверберационный сигнал считается шумом, или оценки второго показателя (108) громкости так, что 40 фильтрованный реверберационный сигнал считается возбуждающим сигналом, а фильтрованный прямой сигнал считается шумом.

3. Устройство по п.1, в котором блок (104b) оценки громкости выполнен с возможностью вычисления первого показателя громкости в качестве громкости фильтрованного прямого сигнала или для вычисления второго показателя громкости 45 в качестве громкости фильтрованного реверберационного сигнала или смешанного сигнала.

4. Устройство по п.1, в котором блок (110) объединения выполнен с возможностью вычисления разности с использованием первого показателя (106) громкости и второго

показателя (108) громкости.

5. Устройство по п.1, дополнительно содержащее:

5 блок (114) предсказания для предсказания воспринимаемого уровня реверберации на основе среднего значения (904), по меньшей мере, двух показателей для воспринимаемой громкости для разных кадров (k) сигнала.

6. Устройство по п.5, в котором блок (114) предсказания выполнен с возможностью использования в предсказании (900) свободного члена (901, 903), линейного члена в зависимости от среднего значения (904) и коэффициента масштабирования (902).

10 7. Устройство по п.5, в котором свободный член (903) зависит от параметра реверберации, описывающего фильтр (606) реверберации, использованный для генерации реверберационного сигнала в искусственном ревербераторе.

8. Устройство по п.1, в котором каскад фильтра содержит каскад (300) частотно-временного преобразования,

15 причем блок (104b) оценки громкости выполнен с возможностью суммирования (104с, 104d) результатов, полученных для множества полос частот, для получения первого и второго показателей (106, 108) громкости для широкополосного смешанного сигнала, содержащего прямую компоненту сигнала и реверберационную компоненту сигнала.

9. Устройство по п.1, в котором каскад (104а) фильтра содержит:

20 фильтр (302) передачи уха, вычислитель (304) модели возбуждения, и временной интегратор (306) для получения фильтрованного прямого сигнала или фильтрованного реверберационного сигнала или фильтрованного смешанного сигнала.

25 10. Способ определения показателя для воспринимаемого уровня реверберации в смешанном сигнале, состоящем из прямой компоненты (100) сигнала и реверберационной компоненты (102) сигнала, содержащий:

30 фильтрацию (104) «сухой» компоненты (100) сигнала, реверберационной компоненты (102) сигнала или смешанного сигнала, причем фильтрацию выполняют с использованием каскада перцепционного фильтра, сконфигурированного для моделирования механизма слухового восприятия объекта для получения фильтрованного прямого сигнала, фильтрованного реверберационного сигнала или фильтрованного смешанного сигнала;

оценку первого показателя громкости с использованием фильтрованного прямого сигнала;

35 оценку второго показателя громкости с использованием фильтрованного реверберационного сигнала или фильтрованного смешанного сигнала, где фильтрованный смешанный сигнал получают из суперпозиции прямой компоненты сигнала и реверберационной компоненты сигнала; и

объединение (110) первого и второго показателей (106, 108) громкости, чтобы получить показатель (112) для воспринимаемого уровня реверберации.

40 11. Аудио процессор для генерации реверберированного сигнала (808) из прямой компоненты (800) сигнала, содержащий:

ревербератор (801) для реверберации прямой компоненты (800) сигнала, чтобы получить реверберированную компоненту (806) сигнала;

45 устройство для определения показателя воспринимаемого уровня реверберации в реверберированном сигнале, содержащем прямую компоненту сигнала и реверберированную компоненту сигнала, в соответствии с одним из пп.1-9;

контроллер (803) для приема воспринимаемого уровня (R), сгенерированного устройством (802) для определения показателя воспринимаемого уровня реверберации,

и для генерации управляющего сигнала (804) в соответствии с воспринимаемым уровнем и целевым значением;

манипулятор (805) для обработки «сухой» компоненты (800) сигнала или реверберационной компоненты (806) сигнала в соответствии с управляющим значением (804); и

блок (807) объединения для объединения обработанной «сухой» компоненты сигнала и обработанной реверберационной компоненты сигнала, или для объединения «сухой» компоненты сигнала и обработанной реверберационной компоненты сигнала, или для объединения обработанной «сухой» компоненты сигнала и реверберационной компоненты сигнала, чтобы получить смешанный сигнал (808).

12. Аудио процессор по п.11, в котором манипулятор (805) содержит блок весовой обработки для взвешивания реверберационной компоненты сигнала значением усиления, причем значение усиления определяется посредством управляющего сигнала, или в котором ревербератор (801) содержит переменный фильтр, причем фильтр способен изменяться в ответ на управляющий сигнал (804).

13. Аудио процессор по п.12, в котором ревербератор (801) имеет фиксированный фильтр,

в котором манипулятор (805) имеет блок весовой обработки, генерирующий обработанную реверберационную компоненту сигнала, и

в котором сумматор (807) сконфигурирован для суммирования прямой компоненты сигнала и обработанной реверберационной компоненты сигнала, чтобы получить смешанный сигнал (808).

14. Способ обработки аудиосигнала для генерации реверберированного сигнала (808) из прямой компоненты (800) сигнала, содержащий:

реверберацию (801) прямой компоненты (800) сигнала, чтобы получить реверберированную компоненту (806) сигнала;

способ определения показателя воспринимаемого уровня реверберации в реверберированном сигнале, содержащем прямую компоненту сигнала и реверберированную компоненту сигнала в соответствии с п.10;

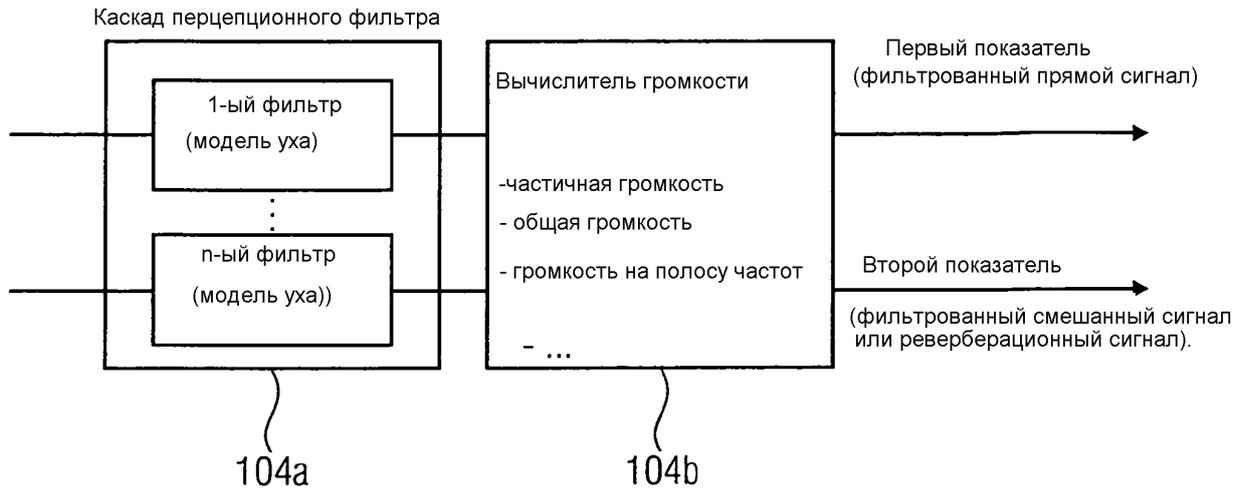
получение воспринимаемого уровня (R), сгенерированного способом (802) для определения показателя воспринимаемого уровня реверберации, генерацию (803) управляющего сигнала (804) в соответствии с воспринимаемым уровнем и целевым значением;

обработку (805) «сухой» компоненты (800) сигнала или реверберационной компоненты (806) сигнала в соответствии с управляющим значением (804); и

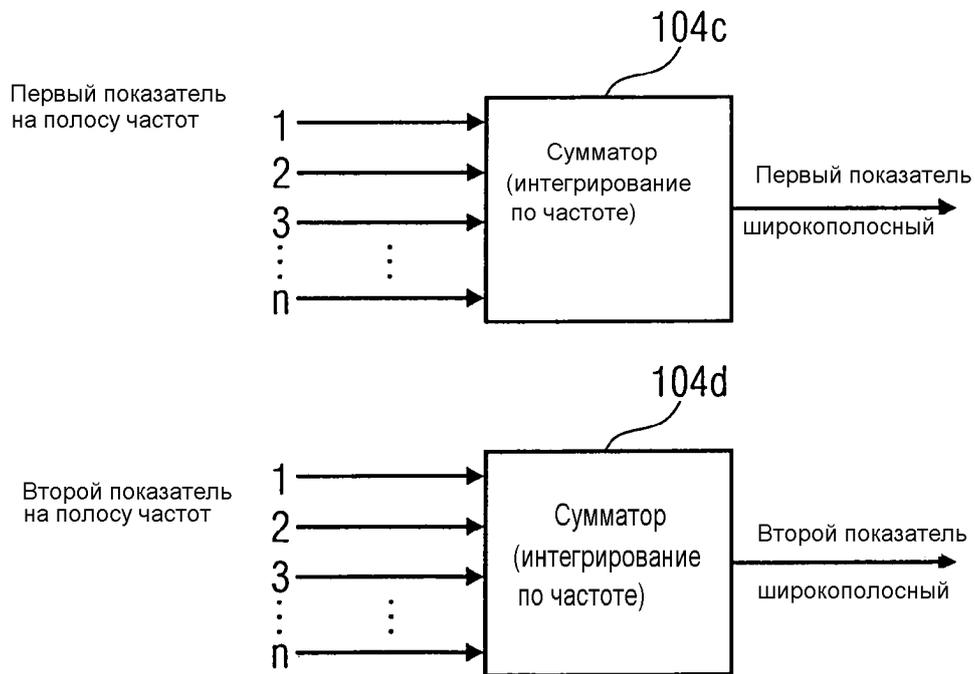
объединение (807) обработанной «сухой» компоненты сигнала и обработанной реверберационной компоненты сигнала, или объединение «сухой» компоненты сигнала и обработанной реверберационной компоненты сигнала, или объединение обработанной «сухой» компоненты сигнала и реверберационной компоненты сигнала, чтобы получить смешанный сигнал (808).

15. Машиночитаемый носитель, содержащий записанную на нем компьютерную программу, имеющую программный код для выполнения, при исполнении на компьютере, способа по п.10.

16. Машиночитаемый носитель, содержащий записанную на нем компьютерную программу, имеющую программный код для выполнения, при исполнении на компьютере, способа по п.14.



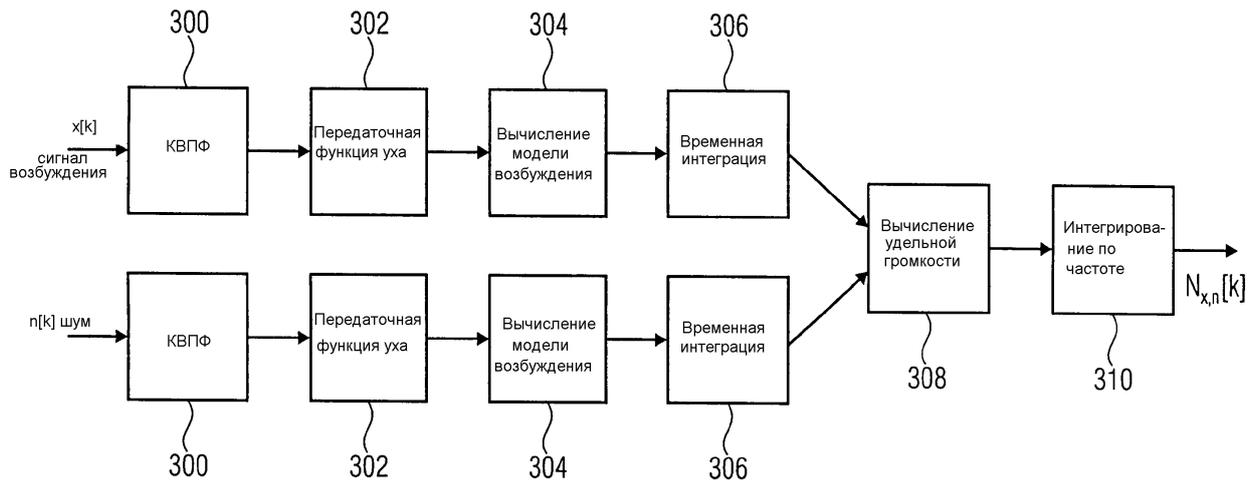
ФИГ.2А



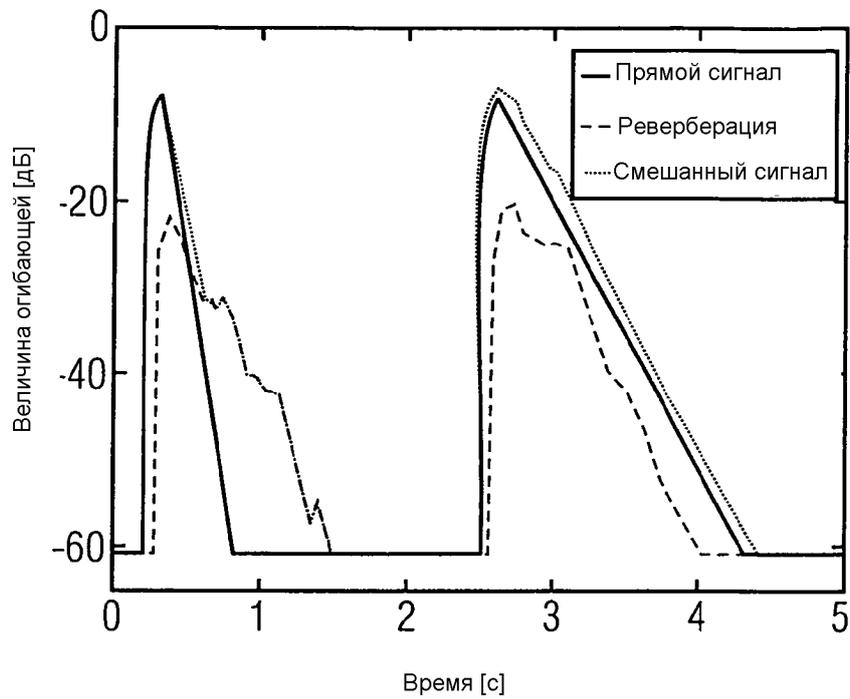
ФИГ.2В

	Модель	EST_1	EST_2	Показатель
1	Частичная	Γ, χ	χ, Γ	$N_{\Gamma, \chi} - N_{\chi, \Gamma}$
2	Общая	m	χ	$N_m - N_\chi$
3	Общая	Γ	m	$N_\Gamma - N_m$
4	Общая	Γ	χ	$N_\Gamma - N_\chi$

ФИГ.2С

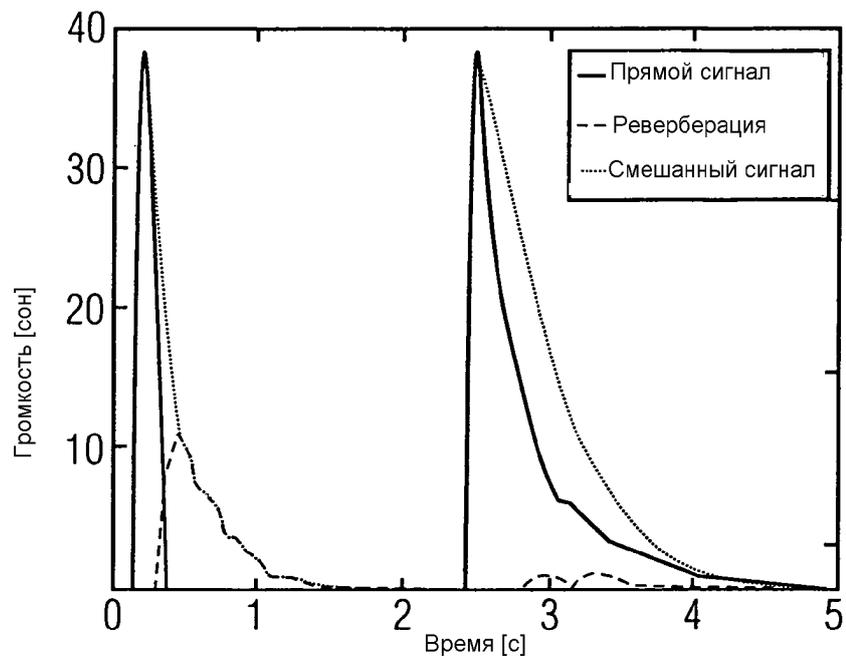


ФИГ.3



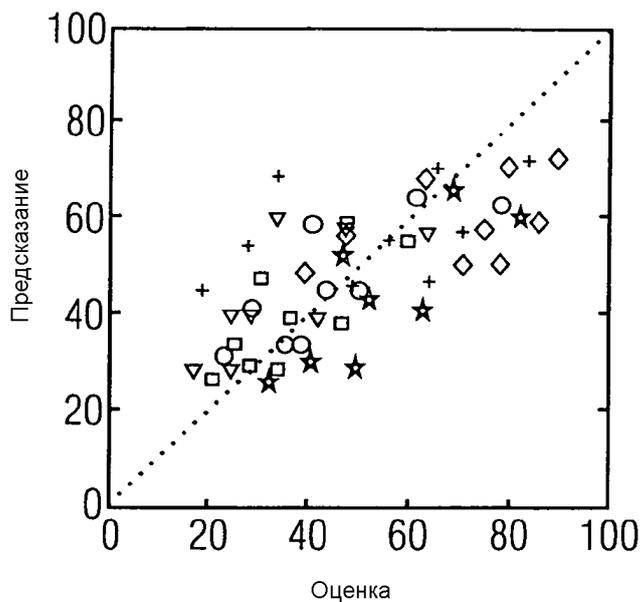
Пример огибающих сигнала-времени аудио сигнала (сплошная линия) реверберационного сигнала (штриховая линия), и смешения обоих сигналов (точечная линия)

ФИГ.4А

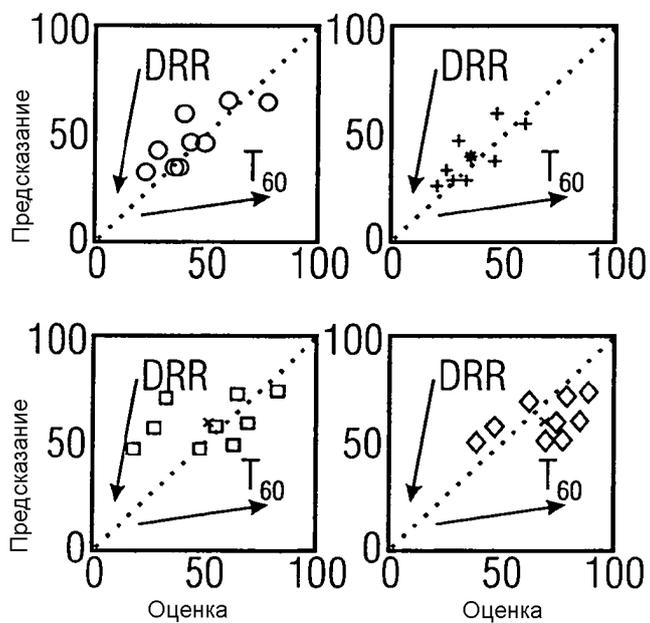


Громкость и частичная громкость примера сигнала: общая громкость (точечная линия), частичная громкость прямого сигнала (сплошная линия) и реверберационного сигнала (штриховая линия),

ФИГ.4В

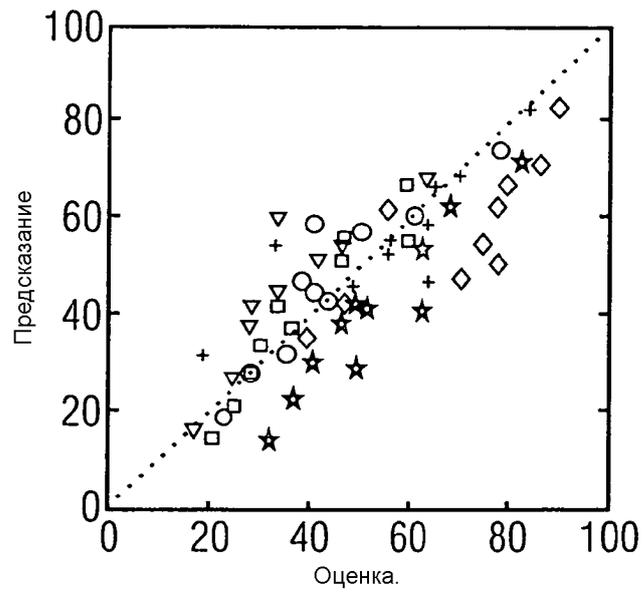


(A) подготовка данных A, все элементы

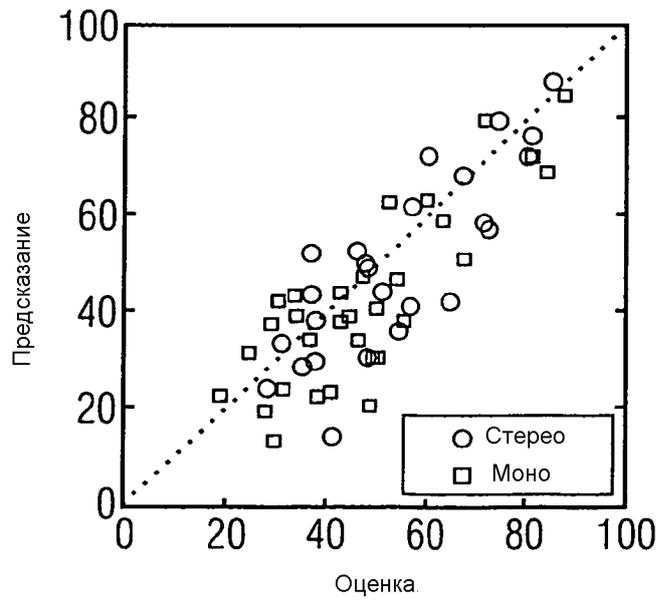


(B) подмножество обучающих данных

ФИГ.5А

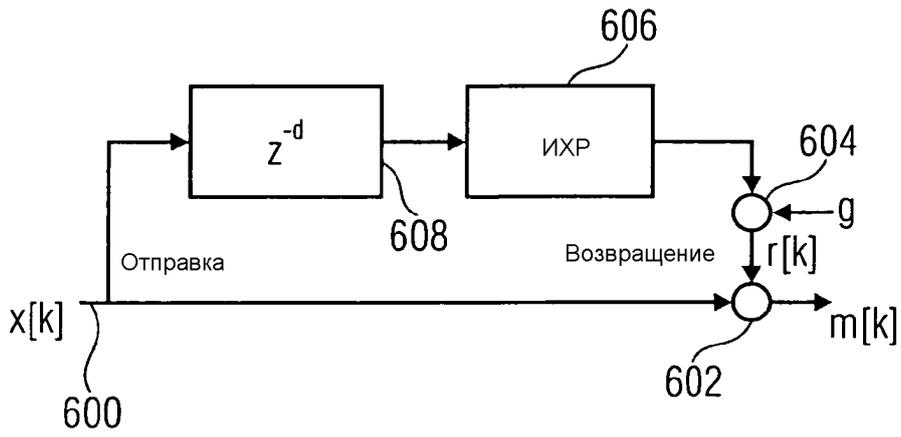


(А) обучающие данные А, символы обозначают прямые сигналы.



(В) тестовые данные В, символы обозначают моно и стерео сигналы

ФИГ.5В



ФИГ.6

	r	САО	СКО
\hat{R}_b, T	0.76	9.7	12.1
\hat{R}_b, V	0.76	10.4	12.6
\hat{R}_e, T	0.85	8.3	10.0
\hat{R}_e, V	0.85	8.3	10.6

Таблица 1. Параметры оценок предсказания \hat{R}_b и \hat{R}_e для обучения (Т) и тестирования (V).



ФИГ.7А

ФИГ.7А	ФИГ.7
ФИГ.7В	

	r	CAO	CKO
$\Delta N_{m-x}, T$	0.81	9.0	11.2
$\Delta N_{m-x}, V$	0.81	9.0	11.1
$\Delta N_{r-m}, T$	0.80	9.5	11.3
$\Delta N_{r-m}, V$	0.80	10.1	12.7
$\Delta N_{r-x}, T$	0.83	8.8	10.5
$\Delta N_{r-x}, V$	0.83	9.1	11.1

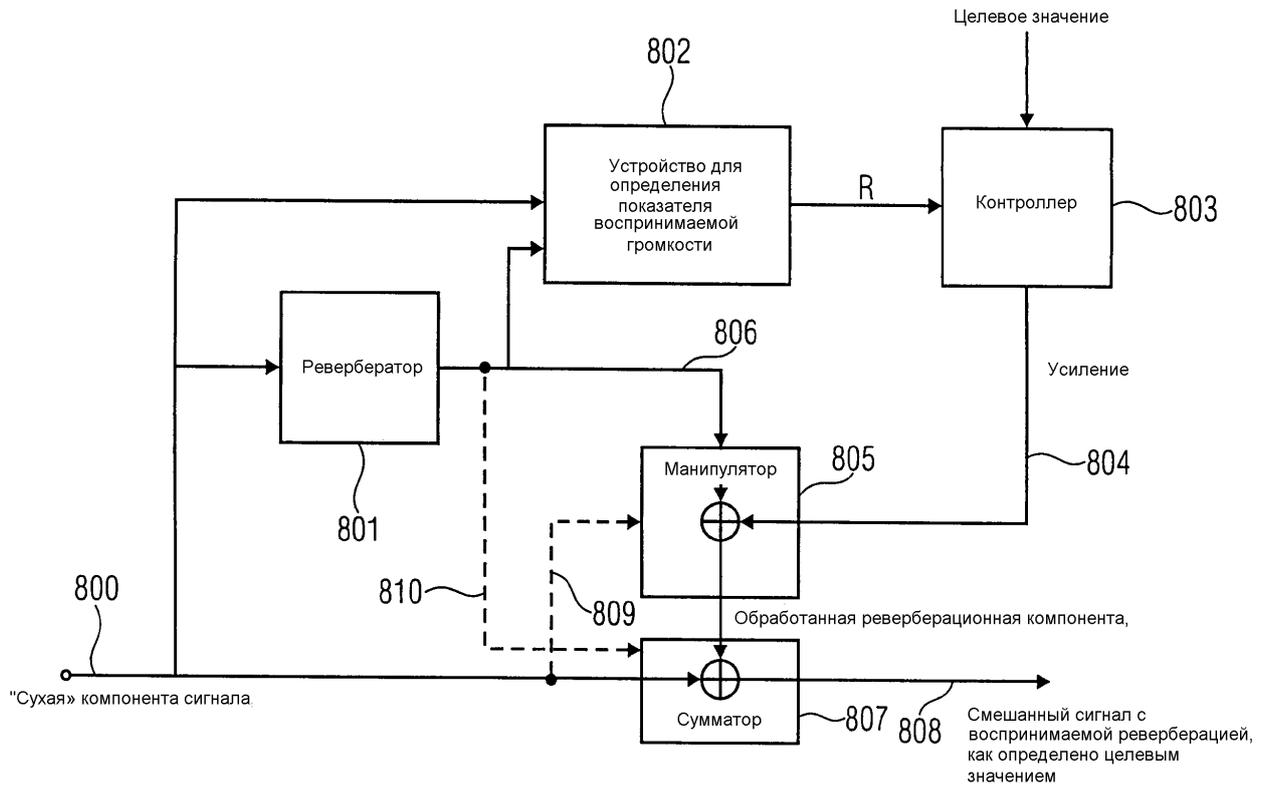
Таблица 2. Параметры оценок, полученные при использовании общей громкости отдельных сигналов и смешанного сигнала, для обучения (Т) и тестирования (V). См. уравнения (5)-(7) для пояснения характеристик громкости

	r	CAO	CKO
Без рассеяния Т	0.84	8.4	10.1
Без рассеяния V	0.84	8.8	10.6
Фиксированное рассеяние Т	0.85	8.3	10.1
Фиксированное рассеяние V	0.85	8.7	10.7

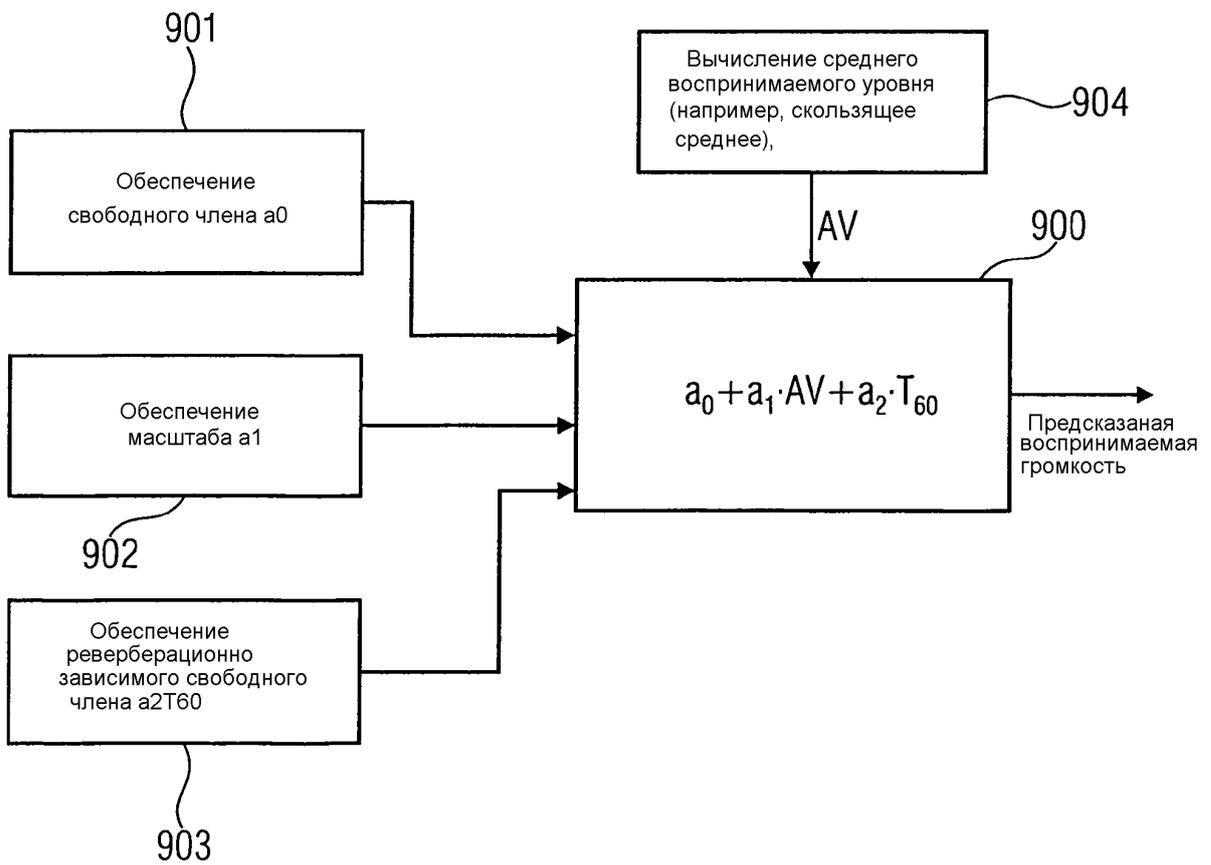
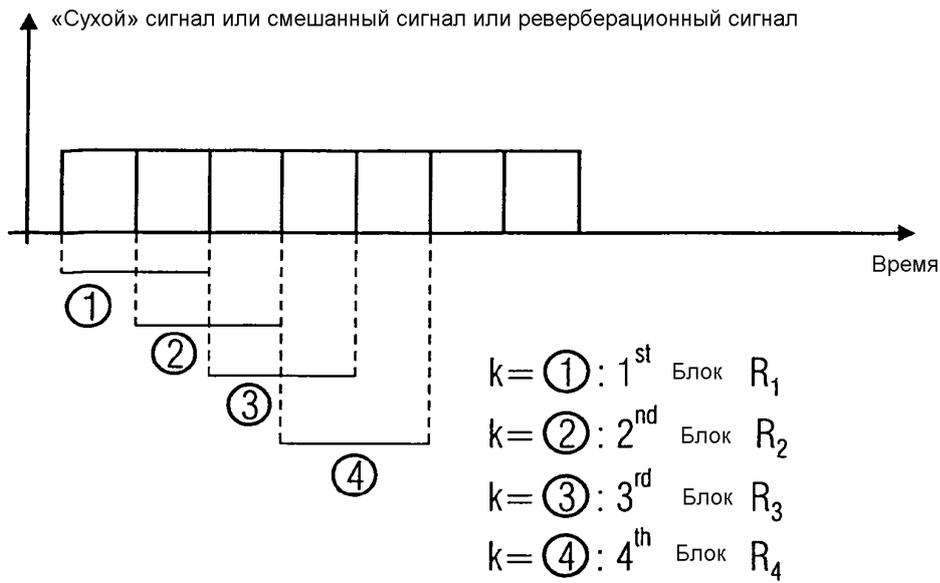
Таблица 3. Параметры оценок для характеристик громкости без рассеяния и с инвариантным по уровням рассеянием, для обучения (Т) и тестирования (V).

ФИГ.7В

ФИГ.7А	ФИГ.7
ФИГ.7В	



ФИГ.8



ФИГ.9

$$\begin{aligned}
 N'_{\text{SIG}} = & C \{ [(E_{\text{SIG}} + E_{\text{NOISE}})G + A]^{\alpha} - A^{\alpha} \} \\
 & - C \{ [(E_{\text{NOISE}}(1+K) + E_{\text{THRQ}})G + A]^{\alpha} \\
 & - (E_{\text{THRQ}}G + A)^{\alpha} \} \left(\frac{E_{\text{THRQ}}}{E_{\text{SIG}}} \right)^{0.3} \quad (17)
 \end{aligned}$$

$$\begin{aligned}
 N'_{\text{SIG}} = & C \left(\frac{2E_{\text{SIG}}}{E_{\text{SIG}} + E_{\text{THRQ}}} \right)^{1.5} \left\{ \frac{(E_{\text{THRQ}}G + A)^{\alpha} - A^{\alpha}}{[(E_{\text{NOISE}}(1+K) + E_{\text{THRQ}})G + A]^{\alpha} - (E_{\text{NOISE}}G + A)^{\alpha}} \right\} \\
 & \times \{ [(E_{\text{SIG}} + E_{\text{NOISE}})G + A]^{\alpha} - (E_{\text{NOISE}}G + A)^{\alpha} \}. \quad (18)
 \end{aligned}$$

$$\begin{aligned}
 N'_{\text{SIG}} = & C_2 (E_{\text{SIG}} + E_{\text{NOISE}})^{0.5} \\
 & - C_2 \{ [(1+K)E_{\text{NOISE}} + E_{\text{THRQ}}]^{0.5} \\
 & - (E_{\text{THRQ}}G + A)^{\alpha} + A^{\alpha} \} \left(\frac{E_{\text{THRQ}}}{E_{\text{SIG}}} \right)^{0.3} \quad (19)
 \end{aligned}$$

$$\begin{aligned}
 N'_{\text{SIG}} = & C \left(\frac{2E_{\text{SIG}}}{E_{\text{SIG}} + E_{\text{THRQ}}} \right)^{1.5} \left\{ \frac{(E_{\text{THRQ}}G + A)^{\alpha} - A^{\alpha}}{[(E_{\text{NOISE}}(1+K) + E_{\text{THRQ}})^{0.5} - (E_{\text{NOISE}})^{0.5}]} \right\} \\
 & \times [(E_{\text{SIG}} + E_{\text{NOISE}})^{0.5} - (E_{\text{NOISE}})^{0.5}]. \quad (20)
 \end{aligned}$$

Уравнения (17) - (20) из В.С.Ж. Moore, В.Р. Glasberg, and Т. Baer, "A model for the prediction of threshold, loudness, and partial loudness," J. Audio Eng. Soc., изд. 45, стр. 224-240, 1997.

ФИГ. 10