



**ФЕДЕРАЛЬНАЯ СЛУЖБА
ПО ИНТЕЛЛЕКТУАЛЬНОЙ СОБСТВЕННОСТИ,
ПАТЕНТАМ И ТОВАРНЫМ ЗНАКАМ**

(12) ОПИСАНИЕ ИЗОБРЕТЕНИЯ К ПАТЕНТУ(21)(22) Заявка: **2008143336/09, 30.03.2007**(24) Дата начала отсчета срока действия патента:
30.03.2007

Приоритет(ы):

(30) Конвенционный приоритет:
04.04.2006 US 60/789,539(43) Дата публикации заявки: **10.05.2010** Бюл. № 13(45) Опубликовано: **10.08.2011** Бюл. № 22(56) Список документов, цитированных в отчете о
поиске: **US 2004190740 A1, 30.09.2004. WO 02/17678**
A1, 28.02.2002. RU 2206174 C2, 20.02.2003.
EP 1366564 B1, 14.02.2002. RU 2012994 C1,
15.05.1994.(85) Дата начала рассмотрения заявки РСТ на
национальной фазе: **05.11.2008**(86) Заявка РСТ:
US 2007/007946 (30.03.2007)(87) Публикация заявки РСТ:
WO 2007/120453 (25.10.2007)

Адрес для переписки:

**129090, Москва, ул. Б.Спасская, 25, стр.3,
ООО "Юридическая фирма Городисский и
Партнеры", пат.пов. Ю.Д.Кузнецову,
рег.№ 595**

(72) Автор(ы):

СИФЕЛЬДТ Алан Джеффри (US)

(73) Патентообладатель(и):

**ДОЛБИ ЛЭБОРЕТЕРИЗ ЛАЙСЕНСИНГ
КОРПОРЕЙШН (US)****(54) РАСЧЕТ И РЕГУЛИРОВКА ВОСПРИНИМАЕМОЙ ГРОМКОСТИ И/ИЛИ
ВОСПРИНИМАЕМОГО СПЕКТРАЛЬНОГО БАЛАНСА ЗВУКОВОГО СИГНАЛА**

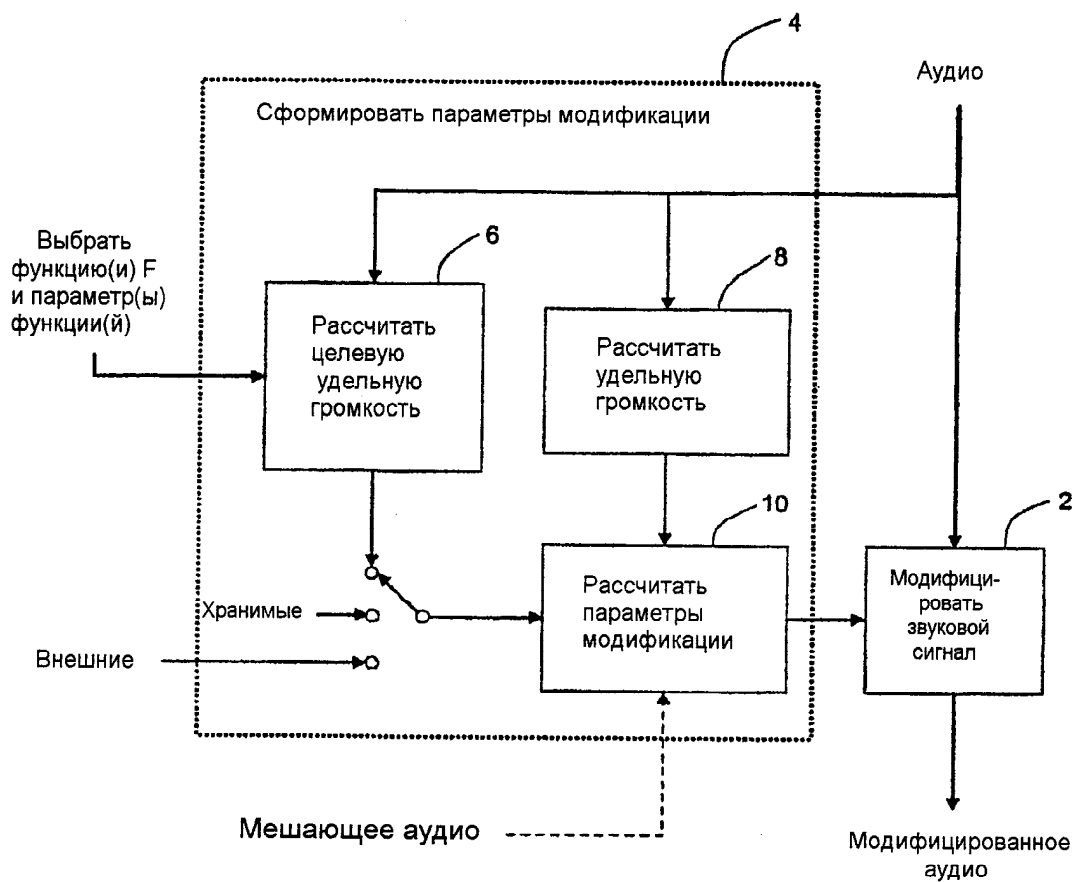
(57) Реферат:

Изобретение относится к обработке звуковых сигналов, относящейся к измерению и регулированию воспринимаемой громкости звука и/или воспринимаемого спектрального баланса звукового сигнала. Обработка звуковых сигналов полезна, например, в одном или более из: регулировки уровня громкости с компенсацией громкости, автоматической регулировки усиления, регулировки

динамического диапазона (в том числе, например, ограничителях, компрессорах, расширителях динамического диапазона и т.п.), динамической коррекции и компенсации шумовых фоновых помех в средах воспроизведения аудио. В различных вариантах осуществления параметры модификации получают для модифицирования звукового сигнала, для того чтобы уменьшать разность между его удельной

громкостью и целевой удельной громкостью.
Технический результат - повышение

разборчивости звукового сигнала. 4 н. и 22 з.п.
ф-лы, 19 ил.



ФИГ. 1



FEDERAL SERVICE
FOR INTELLECTUAL PROPERTY,
PATENTS AND TRADEMARKS

(51) Int. Cl.

G10L 21/02 (2006.01)*H03L 9/00* (2006.01)**(12) ABSTRACT OF INVENTION**(21)(22) Application: **2008143336/09, 30.03.2007**(24) Effective date for property rights:
30.03.2007

Priority:

(30) Priority:
04.04.2006 US 60/789,539(43) Application published: **10.05.2010 Bull. 13**(45) Date of publication: **10.08.2011 Bull. 22**(85) Commencement of national phase: **05.11.2008**(86) PCT application:
US 2007/007946 (30.03.2007)(87) PCT publication:
WO 2007/120453 (25.10.2007)

Mail address:

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

(72) Inventor(s):

SIFEL'DT Alan Dzheffri (US)

(73) Proprietor(s):

**DOLBI LEHBORETERIZ LAJSENSING
KORPOREJShN (US)**

(54) CALCULATION AND ADJUSTMENT OF AUDIO SIGNAL AUDIBLE VOLUME AND/OR SPECTRAL BALANCE

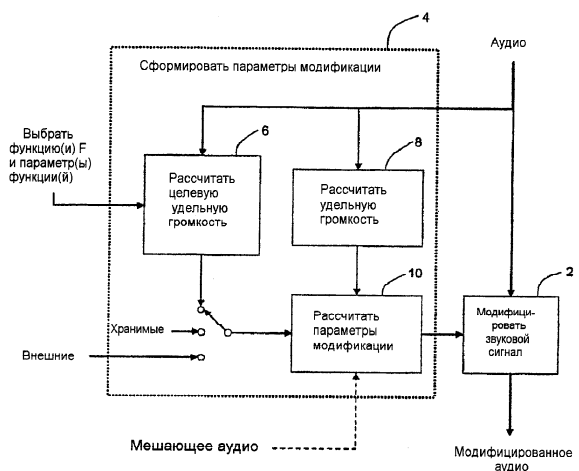
(57) Abstract:

FIELD: physics.

SUBSTANCE: audio signal processing makes sense in volume adjustment with volume compensation, dynamic correction and compensation of background noise in audio playback hardware. Modification parameters are used to modify audio signal to reduce difference between specific volume and target specific volume.

EFFECT: higher intelligibility.

26 cl, 19 dwg



ФИГ. 1

ОБЛАСТЬ ТЕХНИКИ

Изобретение относится к обработке звуковых сигналов. Более точно, изобретение относится к измерению и регулировке воспринимаемой громкости звука и/или воспринимаемого спектрального баланса звукового сигнала. Изобретение, например, полезно в одном или более из: регулировки уровня громкости с компенсацией громкости, автоматической регулировки усиления, регулировки динамического диапазона (в том числе, например, ограничителях, компрессорах, расширителях динамического диапазона и т.п.), динамической коррекции и компенсации шумовых фоновых помех в средах воспроизведения аудио. Изобретение включает в себя не только способы, но также и соответствующие компьютерные программы и устройство.

УРОВЕНЬ ТЕХНИКИ

Было много попыток разработать удовлетворительный объективный способ измерения громкости. Флетчер и Мунсон определили в 1933 году, что человеческий слух менее чувствителен на низких и высоких частотах, чем на средних (или голосовых) частотах. Они также обнаружили, что относительное изменение чувствительности уменьшалось по мере того, как уровень звука увеличивался.

Предыдущий измеритель громкости состоял из микрофона, усилителя, измерителя и соединения фильтров, сконструированных, чтобы грубо копировать частотную характеристику слуха на низких, средних и высоких уровнях звука.

Даже если такие устройства обеспечивали измерения громкости одиночного изолированного тона постоянного уровня, измерения более сложных звуков не очень хорошо соответствовали субъективным ощущениям громкости. Измерители уровня звука этого типа были стандартизованы, но использовались только для специфических задач, таких как дозиметрический контроль и надзор за промышленными шумами.

В начале 1950-х Звикер и Стивенс, среди прочего, продолжили работу Флетчера и Мунсона по разработке более реалистичной модели процесса восприятия громкости. Стивенс опубликовал способ для «Расчета громкости смешанного шума» в журнале Акустического общества Америки в 1956 году, а Звикер опубликовал свою статью «Psychological and Methodical Basis of Loudness» («Психологическая и методическая основа громкости») в Acoustica в 1958 году. В 1959 году Звикер опубликовал графический метод для расчета громкости, а также несколько подобных статей вскоре после этого. Способы Стивенса и Звикера были стандартизованы в качестве ISO 532, частей А и В (соответственно). Оба способа заключали в себе сходные этапы.

Прежде всего, зависящее от времени распределение энергии вдоль базилярной мембраны внутреннего уха, указываемое ссылкой как накачка, имитируется прохождением звукового сигнала через гребенку полосовых слуховых фильтров с центральными частотами, равномерно разнесенными по ступенчатой шкале критических полос. Каждый слуховой фильтр предназначен для имитации частотной характеристики в конкретном местоположении вдоль базилярной мембраны внутреннего уха, с центральной частотой фильтра, соответствующей этому местоположению. Ширина критической полосы определена как ширина полосы пропускания одного такого фильтра. Измеряемая в единицах Герц, ширина критической полосы этих слуховых фильтров увеличивается с увеличением центральной частоты. Поэтому полезно определять криволинейную шкалу частот из условия, чтобы ширина критической полосы для всех слуховых фильтров, измеренная по этой криволинейной шкале, была постоянной. Такая криволинейная шкала указывается ссылкой как ступенчатая шкала критических полос и очень полезна в

понимании и имитации широкого диапазона физиологических феноменов. Например, смотрите *Psychoacoustics - Facts and Models* by E. Zwicker and H. Fastl, Springer-Verlag, Berlin, 1990 (*Психоакустика - факты и модели* по Е. Звикеру и Х.Фасти, Спрингер-Верлаг, Берлин, 1990 год). Способы Стивенса и Звикера используют ступенчатую

5 шкалу критических полос, указываемую ссылкой как шкала Барка, в которой ширина критической полосы является постоянной ниже 500 Гц и увеличивается выше 500 Гц. Позднее, Мур и Глазберг определили ступенчатую шкалу критических полос, которую они называли шкалой, эквивалентной прямоугольной полосы пропускания (ERB) (В. С.

10 J. Moore, B. Glasberg, T. Baer, «A Model for the Prediction of Thresholds, Loudness, and Partial Loudness», *Journal of the Audio Engineering Society*, Vol. 45, No. 4, April 1997, pp. 224-240 (Б. Ц. Дж. Мур, Б. Глазберг, Т. Баер, «Модель для предсказания пороговых значений, громкости и громкости частичных тонов», Журнал сообщества звукотехники, том 45, № 4, Апрель 1997 г., стр. 224-240)). Благодаря психоакустическим экспериментам с

15 использованием маскеров шума с узкополосным провалом в спектре Мур и Глазберг продемонстрировали, что ширина критической полосы продолжает уменьшаться ниже 500 Гц, в противоположность шкале Барка, где ширина критической полосы остается постоянной.

20 Последующее вычисление накачки является функцией нелинейного сжатия, которая формирует параметр указываемый ссылкой как «удельная громкость». Удельная громкость является мерой громкости восприятия в качестве функции частоты и времени и может измеряться в единицах громкости восприятия на единичную частоту по ступенчатой шкале критических полос, такой как шкала Барка или ERB,

25 обсужденная выше. Умозрительно, удельная громкость представляет непрерывное распределение громкости в качестве функции частоты и времени, а зависящая от времени «полная громкость» вычисляется интегрированием этого распределения по частоте. На практике точное восприятие удельной громкости получается

30 дискретизацией этого распределения равномерно по ступенчатой шкале критических полос, например, посредством использования слуховых фильтров, упомянутых выше. В этом случае полная громкость может вычисляться простым суммированием удельной громкости из каждого фильтра. Для уменьшения сложности некоторые приложения могут вычислять грубое приближение для удельной громкости за счет

35 незначительных неточностей в оценке и модификации воспринимаемой громкости. Такие приближения позже будут обсуждены более подробно.

Громкость может измеряться в единицах фонов. Громкостью заданного в фонах звука является уровень звукового давления (SPL) тона в 1 кГц, имеющий

40 субъективную громкость, равную таковой у звука. Традиционно началом отсчета 0 дБ для SPL является среднеквадратическое давление 2×10^{-5} Паскалей и поэтому это также является началом отсчета 0 фонов. Используя это определение при сравнении громкости тонов на частотах, иных чем 1 кГц, с громкостью на 1 КГц, может быть

45 определена кривая равной громкости для заданного в фонах уровня. Фиг.11 показывает кривые равной громкости для частот между 20 Гц и 12,5 кГц, и для уровней в фонах между 4,2 фона (считается порогом слышимости) и 120 фонами (ISO226: 1087 (E), «Acoustics - Normal equal loudness level contours» («Акустика - нормальные кривые равного уровня громкости»)). Измерение в фонах учитывает

50 меняющуюся чувствительность человеческого слуха в зависимости от частоты, но результаты не предоставляют возможности оценки относительных субъективных громкостей звука при переменных уровнях, так как нет попытки ввести поправку на нелинейность увеличения громкости в зависимости от SPL, то есть на то

обстоятельство, что интервал кривых меняется.

Громкость также может измеряться в единицах «сонов». Есть однозначное соответствие между единицами фонов и единицам сонов, которое указано на фиг.11. Один сон определен в качестве громкости немодулированной гармонической волны 1 кГц при 40 дБ (SPL) и равен 40 фонам. Единицы сонов являются такими, что двойное увеличение в сонах соответствует удвоению воспринимаемой громкости. Например, 4 сона воспринимаются как громкость, вдвое большая той, что в 2 сона. Таким образом, выражение уровней громкости в сонах является более информативным. При условии определения удельной громкости как показателя громкости восприятия в качестве функции частоты и времени удельная громкость может измеряться в единицах сонов на единичную частоту. Таким образом, при использовании шкалы Барка удельная громкость обладает единицами сонов на Барк и, подобным образом, с использованием шкалы ERB единицами являются соны на ERB.

Как упомянуто выше, чувствительность человеческого уха изменяется как в зависимости от частоты, так и от уровня, обстоятельство, хорошо документированное в литературе по психоакустике. Одно из следствий состоит в том, что воспринимаемый спектр или тембр данного звука меняется в зависимости от акустического уровня, при котором звук прослушивается. Например, для звука, содержащего низкие, средние и высокие частоты, воспринимаемые относительные пропорции таких частотных составляющих изменяются с общей громкостью звука; когда она тихая, низкие и высокие частотные составляющие звучат тише относительно средних частот, чем они звучат, когда она громкая. Это явление общеизвестно, и было уменьшено в оборудовании воспроизведения звука посредством так называемых тонкомпенсированных регуляторов громкости. Тонкомпенсированный регулятор громкости является регулятором уровня громкости, который применяет низкочастотный, а иногда также и высокочастотный подъем по мере того как уровень громкости убавляется. Таким образом, меньшая чувствительность уха на крайних значениях частот компенсируется искусственным подъемом таких частот. Такие регуляторы являются полностью пассивными; степень применяемой компенсации является функцией настройки регулятора громкости или некоторого другого управляемого пользователем регулятора, не в качестве функции контента звуковых сигналов.

На практике изменения воспринимаемого относительного спектрального баланса между низкими, средними и высокими частотами зависят от сигнала, в частности от его действующего спектра и от того, предназначено ли ему быть громким или тихим. Рассмотрим запись симфонического оркестра. Воспроизводимый на одном и том же уровне, который слышал бы член публики в концертном зале, баланс по ширине спектра может быть правильным, громко или тихо играет оркестр. Если музыка воспроизводится, например, тише на 10 дБ, воспринимаемый баланс по ширине спектра изменяется одним образом для громких пассажей и изменяется другим образом для тихих пассажей. Традиционный пассивный тонкомпенсированный регулятор громкости не применяет разные компенсации в качестве функции музыки.

В международной патентной заявке № PCT/US 2004/016964, зарегистрированной 27 мая 2004 года, опубликованной 23 декабря 2004 года в качестве WO 2004/111994 A2, Шифельдт и другие раскрывают, среди прочего, систему для измерения и настройки воспринимаемой громкости звукового сигнала. Упомянутая заявка PCT, которая указывает Соединенные Штаты, настоящим включена в состав посредством ссылки во всей своей полноте. В упомянутой заявке психоакустическая модель рассчитывает

громкость звукового сигнала в единицах восприятия. В дополнение заявка учреждает технологию для вычисления широкополосного мультипликативного коэффициента усиления, который, когда применяется к аудио, дает в результате громкость модифицированного по коэффициенту усиления аудио, по существу, являющуюся такой же, как эталонная громкость. Однако применение такого широкополосного усиления изменяет воспринимаемый спектральный баланс аудио.

РАСКРЫТИЕ ИЗОБРЕТЕНИЯ

В одном из аспектов изобретение предусматривает извлечение информации, используемой для регулирования удельной громкости звукового сигнала модифицированием звукового сигнала, для того чтобы уменьшить разницу между его удельной громкостью и целевой удельной громкостью. Удельная громкость является мерой громкости восприятия в качестве функции частоты и времени. В практических реализациях удельная громкость модифицированного звукового сигнала может делаться приближающейся к целевой удельной громкости. Приближение может находиться под влиянием не только соображений обычной сигнальной обработки, но также и временного и/или частотного сглаживания, которое может применяться при модифицировании, как описано ниже.

Так как удельная громкость является мерой громкости восприятия звукового сигнала как функции частоты и времени, для того чтобы уменьшить разность между удельной громкостью звукового сигнала и целевой удельной громкостью, модифицирование может модифицировать звуковой сигнал в качестве функции частоты. Хотя в некоторых случаях целевая удельная громкость может быть не зависящей от времени, и сам звуковой сигнал может быть установившимся не зависящим от времени сигналом, типично модифицирование также может модифицировать звуковой сигнал в качестве функции времени.

Аспекты настоящего изобретения также могут применяться для компенсации фонового шума, вмешивающегося в среду воспроизведения аудио. Когда аудио прослушивается в присутствии фонового шума, шум может частично или полностью маскировать аудио некоторым образом, зависимым как от уровня и спектра аудио, так и от уровня и спектра шума. Результатом является перестройка воспринимаемого спектра аудио. В соответствии с психоакустическим учением (например, смотрите Moore, Glasberg, and Baer, «A Model for the Prediction of Thresholds, Loudness, and Partial Loudness», *J. Audio Eng. Soc.*, Vol. 45, No. 4, April 1997 (Мур, Глазберг и Баер, «Модель для предсказания пороговых значений, громкости и громкости частичных тонов», журнал сообщества звукотехники, том 45, №4, апрель 1997 г.)), можно определять «удельную громкость частичных тонов» аудио как громкость восприятия аудио в присутствии вторичного мешающего звукового сигнала, такого как шум.

Таким образом, в еще одном аспекте изобретение предусматривает извлечение информации, используемой для регулирования удельной громкости звукового сигнала, модифицированием звукового сигнала, для того чтобы уменьшить разницу между его удельной громкостью частичных тонов и целевой удельной громкостью. Выполнение этого смягчает влияния шума точным по ощущениям образом. В этом и других аспектах изобретения, которые учитывают мешающий шумовой сигнал, предполагается, что есть доступ отдельно к звуковому сигналу и отдельно к вторичному мешающему сигналу.

В еще одном аспекте изобретение предусматривает регулирование удельной громкости звукового сигнала модифицированием звукового сигнала, для того чтобы уменьшать разницу между его удельной громкостью и целевой удельной громкостью.

В еще одном аспекте изобретение предусматривает регулирование удельной громкости частичных тонов звукового сигнала модифицированием звукового сигнала, для того чтобы уменьшать разницу между его удельной громкостью и целевой удельной громкостью.

Когда целевая удельная громкость не является функцией звукового сигнала, она может быть хранимой и принимаемой целевой удельной громкостью. Когда целевая удельная громкость не является функцией звукового сигнала, модифицирование или получение может явно или неявно рассчитывать удельную громкость или удельную громкость частичных тонов. Примеры неявного расчета включают в себя справочную таблицу или «отражающее ряд решений» математическое выражение, в котором удельная громкость и/или удельная громкость частичных тонов определяется по своей природе (термин, отражающий ряд решений, упомянут для описания математического выражения, которое может быть точно представлено с использованием конечного количества стандартных математических операций и функций, таких как возведение в степень и косинус). К тому же, когда целевая удельная громкость не является функцией звукового сигнала, целевая удельная громкость может быть независимой как от времени, так и от частоты, или она может быть независимой только от времени.

В еще одном другом аспекте изобретение предусматривает обработку звукового сигнала посредством обработки звукового сигнала или показателя звукового сигнала в соответствии с одной или более последовательностей операций или одним или более параметрами управления последовательностью операций для формирования целевой удельной громкости. Хотя целевая удельная громкость может быть независимой от времени («неизменной»), целевая удельная громкость преимущественно может быть функцией удельной громкости звукового сигнала. Хотя она может быть статическим, независимым от частоты и времени сигналом, типично сам звуковой сигнал является зависящим от частоты и времени, таким образом заставляя целевую удельную громкость быть зависящей от частоты и времени, когда она является функцией звукового сигнала.

Аудио и целевая удельная громкость или представление целевой удельной громкости могут приниматься из передаваемых данных или воспроизводиться с запоминающего носителя.

Представление целевой удельной громкости может быть одним или более масштабными коэффициентами, которые масштабируют звуковой сигнал или показатель звукового сигнала.

Целевая удельная громкость любого из вышеприведенных аспектов изобретения может быть функцией звукового сигнала или показателя звукового сигнала. Одним из подходящих показателей звукового сигнала является удельная громкость звукового сигнала. Функция звукового сигнала или показателя звукового сигнала может быть масштабированием звукового сигнала или показателя звукового сигнала. Например, масштабирование может быть одним или комбинацией из масштабирований:

(а) зависящего от времени и частоты масштабного коэффициента $\Xi[b, t]$, масштабирующего удельную громкость, как в зависимости

$$\hat{N}[b, t] = \Xi[b, t]N[b, t];$$

(b) зависящего от времени, независящего от частоты масштабного коэффициента $\Phi[t]$, масштабирующего удельную громкость, как в зависимости

$$\hat{N}[b, t] = \Phi[t]N[b, t];$$

(с) независящего от времени, зависящего от частоты масштабного коэффициента $\Theta[b]$, масштабирующего удельную громкость, как в зависимости

$$\hat{N}[b, t] = \Theta[b]N[b, t]; \text{ и}$$

5 (d) независящего от времени, независящего от частоты масштабного коэффициента α , масштабирующего удельную громкость звукового сигнала, как в зависимости

$$\hat{N}[b, t] = \alpha N[b, t],$$

10 в которых $\hat{N}[b, t]$ - целевая удельная громкость, $M[b, t]$ - удельная громкость звукового сигнала, b - показатель частоты, а t - показатель времени.

В случае (а) зависящего от времени и частоты масштабного коэффициента масштабирование может определяться по меньшей мере частично отношением требуемой многополосной громкости и многополосной громкости звукового сигнала. Такое масштабирование может быть используемым в качестве регулятора динамического диапазона. Дополнительные подробности аспектов применения изобретения в качестве регулятора динамического диапазона изложены ниже.

15 К тому же в случае (а) зависящего от времени и частоты масштабного коэффициента удельная громкость может масштабироваться отношением показателя требуемой спектральной формы к показателю спектральной формы звукового сигнала. Такое масштабирование может применяться для преобразования воспринимаемого спектра звукового сигнала из зависящего от времени воспринимаемого спектра в по существу независящий от времени воспринимаемый спектр. Когда удельная громкость масштабируется отношением показателя требуемой спектральной формы к показателю спектральной формы звукового сигнала, такое масштабирование может быть используемым в качестве динамического эквалайзера. Дополнительные подробности аспектов применения изобретения в качестве динамического эквалайзера изложены ниже.

30 В случае (b) зависящего от времени, независящего от частоты масштабного коэффициента, масштабирование может определяться по меньшей мере частично отношением требуемой широкополосной громкости и широкополосной громкости звукового сигнала. Такое масштабирование может быть используемым в качестве автоматического регулятора усиления или регулятора динамического диапазона. 35 Дополнительные подробности аспектов применения изобретения в качестве автоматического регулятора усиления и регулятора динамического диапазона изложены ниже.

40 В случае (а) (зависящего от времени и частоты масштабного коэффициента) или случая (b) (зависящего от времени, независящего от частоты масштабного коэффициента) масштабный коэффициент может быть функцией звукового сигнала или показателем звукового сигнала.

В обоих, случае (с) независящего от времени, зависящего от частоты масштабного коэффициента или случае (d) зависящего от времени, независящего от частоты масштабного коэффициента, модифицирование или получение может включать в себя хранение масштабного коэффициента, или масштабный коэффициент может 45 приниматься из внешнего источника.

В любом из случаев (с) и (d) масштабный коэффициент может не быть функцией звукового сигнала или показателя звукового сигнала.

50 В любом из различных аспектов изобретения и его вариантов модифицирование, получение или формирование могут по-разному явно или неявно рассчитывать (1) удельную громкость и/или (2) удельную громкость частичных тонов, и/или (3) целевую

удельную громкость. Неявные расчеты, например, могут заключать в себе справочную таблицу или отражающее ряд решений математическое выражение.

Параметры модификации могут быть сглаженными во времени. Параметрами модификации, например, могут быть (1) множество коэффициентов масштабирования амплитуды, относящихся к полосам частот звукового сигнала, или (2) множество коэффициентов фильтра для управления одним или более фильтрами, такими как многоотводный (с конечной импульсной характеристикой, FIR) КИХ-фильтр или многополюсный (с бесконечной импульсной характеристикой, IIR) БИХ-фильтр.

Коэффициенты масштабирования или коэффициенты фильтра (и фильтры, к которым они применяются) могут быть зависящими от времени.

При расчете функции удельной громкости звукового сигнала, которая определяет целевую удельную громкость, или инверсии такой функции, последовательность операций или последовательности операций, выполняющие такие расчеты, работают в том, что может быть охарактеризовано как область (психоакустической) громкости восприятия - входными данными и выходными данными расчета являются удельные громкости. В противоположность при применении коэффициентов масштабирования амплитуды к полосам частот звукового сигнала или применении коэффициентов фильтра к регулируемой фильтрации звукового сигнала параметры модификации действуют для модифицирования звукового сигнала вне области (психоакустической) громкости восприятия, в том, что может характеризоваться как область электрических сигналов. Хотя модификации в отношении звукового сигнала могут производиться в отношении звукового сигнала в области электрических сигналов, такие изменения в области электрических сигналов получаются из расчетов в области (психоакустической) громкости восприятия, из условия, чтобы модифицированный звуковой сигнал имел удельную громкость, которая приближается к требуемой целевой удельной громкости.

Получением параметров модификации из расчетов в области громкости может достигаться больший контроль над громкостью восприятия и спектральным балансом восприятия, чем если бы такие параметры модификации получались в области электрических сигналов. В дополнение использование психоакустической гребенки фильтров имитации базилярной мембраны или ее эквивалентов при выполнении расчетов в области громкости может обеспечивать более детальное регулирование воспринимаемого спектра, чем в компоновках, которые получают параметры модификации в области электрических сигналов.

Каждое из модифицирования, получения и формирования может быть зависимым от одного или более из показателя мешающего звукового сигнала, целевой удельной громкости, оценки удельной громкости немодифицированного звукового сигнала, полученной из удельной громкости или удельной громкости частичных тонов модифицированного звукового сигнала, удельной громкости немодифицированного звукового сигнала и приближения к целевой удельной громкости, полученного из удельной громкости или удельной громкости частичных тонов модифицированного звукового сигнала.

Модифицирование или получение могут получать параметры модификации по меньшей мере частично из одного или более из показателя мешающего звукового сигнала, целевой удельной громкости, оценки удельной громкости немодифицированного звукового сигнала, полученной из удельной громкости или удельной громкости частичных тонов модифицированного звукового сигнала, удельной громкости немодифицированного звукового сигнала и приближения к

целевой удельной громкости, полученного из удельной громкости или удельной громкости частичных тонов модифицированного звукового сигнала.

Более точно, модифицирование или получение могут получать параметры модификации по меньшей мере частично из

(1) одного из

целевой удельной громкости, и

оценки удельной громкости немодифицированного звукового сигнала, полученной из удельной громкости модифицированного звукового сигнала, и

(2) одного из

удельной громкости немодифицированного звукового сигнала, и

приближения к целевой удельной громкости, полученного из удельной громкости модифицированного звукового сигнала,

или, когда должен учитываться мешающий звуковой сигнал, модифицирование или получение могут получать параметры модификации по меньшей мере частично из

(1) показателя мешающего звукового сигнала,

(2) одного из

целевой удельной громкости, и

оценки удельной громкости немодифицированного звукового сигнала, полученной из удельной громкости частичных тонов модифицированного звукового сигнала, и

(3) одного из

удельной громкости немодифицированного звукового сигнала, и

приближения к целевой удельной громкости, полученного из удельной громкости частичных тонов модифицированного звукового сигнала.

Может применяться компоновка с прямой связью, в которой удельная громкость получается из звукового сигнала и в которой целевая удельная громкость принимается из источника, внешнего по отношению к способу, или из хранения, когда модифицирование или получение включает в себя хранение целевой удельной громкости. В качестве альтернативы может применяться компоновка со смешанной прямой связью/обратной связью, в которой приближение к целевой удельной громкости получается из модифицированного звукового сигнала, и в которой целевая удельная громкость принимается из источника, внешнего по отношению к способу, или из хранения, когда модифицирование или получение включает в себя хранение целевой удельной громкости.

Модифицирование или получение могут включать в себя одну или более последовательностей операций для получения, явно или неявно, целевой удельной громкости, таковая или таковые из которых рассчитывают, явно или неявно, функцию звукового сигнала или показателя звукового сигнала. В одном из альтернативных вариантов может применяться компоновка с прямой связью, в которой удельная громкость и целевая удельная громкость получаются из звукового сигнала, получение целевой удельной громкости применяет функцию звукового сигнала или показателя звукового сигнала. В одном из альтернативных вариантов, может применяться компоновка со смешанной прямой связью/обратной связью, в которой приближение целевой удельной громкости получается из модифицированного звукового сигнала, а целевая удельная громкость получается из звукового сигнала, получение целевой удельной громкости применяет функцию звукового сигнала или показателя звукового сигнала.

Модифицирование или получение могут включать в себя одну или более последовательностей операций для получения, явно или неявно, оценки удельной

громкости немодифицированного звукового сигнала в ответ на модифицированный звуковой сигнал, таковая или таковые из которых рассчитывают, явно или неявно, инверсию функции звукового сигнала или показателя звукового сигнала. В одном из альтернативных вариантов применяется компоновка с обратной связью, в которой оценка удельной громкости немодифицированного звукового сигнала и приближение к целевой удельной громкости получаются из модифицированного звукового сигнала, оценка удельной громкости рассчитывается с использованием инверсии функции звукового сигнала или показателя звукового сигнала. В еще одном альтернативном варианте применяется компоновка со смешанной прямой связью/обратной связью, в которой удельная громкость получается из звукового сигнала, а оценка удельной громкости немодифицированного звукового сигнала получается из модифицированного звукового сигнала, получение оценки рассчитывается с использованием инверсии упомянутой функции звукового сигнала или показателя звукового сигнала.

Параметры модификации могут применяться к звуковому сигналу для формирования модифицированного звукового сигнала.

Еще один аспект изобретения состоит в том, что может быть временное и/или пространственное разделение последовательностей операций или устройств, так что, в действительности, есть кодировщик или кодирование, а также декодер или декодирование. Например, может быть система кодирования/декодирования, в которой модифицирование или получение может передавать и принимать или хранить, а также воспроизводить звуковой сигнал и либо (1) параметры модификации либо (2) целевую удельную громкость или представление целевой удельной громкости. В качестве альтернативы, в действительности, может быть только кодировщик или кодирование, в котором есть передача или хранение звукового сигнала и (1) параметров модификации, либо (2) целевой удельной громкости или представления целевой удельной громкости. В качестве альтернативы, как упомянуто выше, в действительности, может быть только декодер или декодирование, в котором есть прием или воспроизведение звукового сигнала и (1) параметров модификации либо (2) целевой удельной громкости или представления целевой удельной громкости.

ОПИСАНИЕ ЧЕРТЕЖЕЙ

Фиг.1 - функциональная структурная схема, иллюстрирующая пример реализации с прямой связью согласно аспектам изобретения.

Фиг.2 - функциональная структурная схема, иллюстрирующая пример реализации с обратной связью согласно аспектам изобретения.

Фиг.3 - функциональная структурная схема, иллюстрирующая пример реализации со смешанной прямой связью/обратной связью согласно аспектам изобретения.

Фиг.4 - функциональная структурная схема, иллюстрирующая пример еще одной реализации со смешанной прямой связью/обратной связью согласно аспектам изобретения.

Фиг.5 - функциональная структурная схема, иллюстрирующая образ действий, которым немодифицированный звуковой сигнал и параметры модификации, которые определены любой одной из компоновок с прямой связью, обратной связью или со смешанной прямой связью/обратной связью, могут храниться или передаваться для использования, например, в разделенных временным и пространственным образом устройстве или последовательности операций.

Фиг.6 - функциональная структурная схема, иллюстрирующая образ действий, которым немодифицированный звуковой сигнал и целевая удельная громкость или ее

представление, которые определены любой одной из компоновок с прямой связью, обратной связью или со смешанной прямой связью/обратной связью, могут храниться или передаваться для использования, например, в разделенных временным и пространственным образом устройстве или последовательности операций.

Фиг.7 - схематическая функциональная структурная схема или схематическая блок-схема последовательности операций способа, показывающая общее представление аспекта настоящего изобретения.

Фиг.8 - идеализированная типовая характеристика линейного фильтра $P(z)$, пригодного в качестве фильтра передачи в варианте осуществления настоящего изобретения, в котором вертикальной осью является затухание в децибелах (дБ), а горизонтальной осью является логарифмическая, по основанию 10, частота в Герцах (Гц).

Фиг.9 показывает зависимость между шкалой частот ERB (вертикальная ось) и частотой в Герцах (горизонтальная ось).

Фиг.10 показывает набор идеализированных типовых характеристик слухового фильтра, которые аппроксимируют определение критической полосы по шкале ERB. Горизонтальной шкалой является частота в Герцах, а вертикальной шкалой является уровень в децибелах.

Фиг.11 показывает кривые равной громкости по ISO 226. Горизонтальной шкалой является частота в Герцах (логарифмическая, по основанию 10, шкала), а вертикальной шкалой является уровень звукового давления в децибелах.

Фиг.12 показывает кривые равной громкости по ISO 226, нормализованные фильтром $P(z)$ передачи. Горизонтальной шкалой является частота в Герцах (логарифмическая, по основанию 10, шкала), а вертикальной шкалой является уровень звукового давления в децибелах.

Фиг.13a - идеализированный график, показывающий широкополосные и многополосные коэффициенты усиления для масштабирования громкости в 0,25 на сегменте женской речи. Горизонтальной шкалой являются полосы ERB, а вертикальной шкалой является относительный коэффициент усиления в децибелах (dB).

Фиг.13b - идеализированный график, показывающий удельную громкость соответственно исходного сигнала, модифицированного широкополосным коэффициентом усиления сигнала, и модифицированного многополосным коэффициентом усиления сигнала. Горизонтальной шкалой являются полосы ERB, а вертикальной шкалой является удельная громкость (сон/ERB).

Фиг.14a - идеализированный график, показывающий: $L_o[t]$ в качестве функции $L_i[t]$ для типичной АРУ. Горизонтальной шкалой является $\log(L_i[t])$, а вертикальной шкалой является $\log(L_o[t])$.

Фиг.14b - идеализированный график, показывающий: $L_o[t]$ в качестве функции $L_i[t]$ для типичной DRC. Горизонтальной шкалой является $\log(L_i[t])$, а вертикальной шкалой является $\log(L_o[t])$.

Фиг.15 - идеализированный график, показывающий типичную функцию сглаживания полос для многополосной DRC. Горизонтальной шкалой является номер полосы, а вертикальной шкалой является выход коэффициента усиления для полосы b.

Фиг.16 - схематическая функциональная структурная схема или схематическая блок-схема последовательности операций способа, показывающая общее представление аспекта настоящего изобретения.

Фиг.17 - схематическая функциональная структурная схема или схематическая блок-схема последовательности операций способа, подобные фиг.1, которая к тому же

включает в себя компенсацию шума в среде воспроизведения.

НАИЛУЧШИЙ ВАРИАНТ ОСУЩЕСТВЛЕНИЯ ИЗОБРЕТЕНИЯ

Фиг.с 1 по 4 показывают функциональные структурные схемы, иллюстрирующие возможные примеры реализаций с прямой связью обратной связью и два варианта со смешанной прямой связью/обратной связью, согласно аспектам изобретения.

Со ссылкой на пример топологии с прямой связью на фиг.1 звуковой сигнал подается в два тракта: (1) сигнальный тракт, содержащий последовательность операций, или устройство 2 («Модифицировать звуковой сигнал»), способные к модификации аудио в ответ на параметры модификации, (2) тракт управления, содержащий последовательность управления, или устройство 4 («Сформировать параметры модификации»), способные к формированию таких параметров модификации. Модифицировать звуковой сигнал 2 в примере топологии с прямой связью фиг.1 и в каждом из примеров фиг.2-4 можно устройством или последовательностью операций, которые модифицируют звуковой сигнал, например его амплитуду, зависящий от частоты и/или времени образом в соответствии с параметрами М модификации, принятыми из Сформировать параметры модификации, 4, (или из эквивалентных последовательностей операций или устройств 4', 4" и 4"', в каждом из примеров фиг.2-4 соответственно). Сформировать параметры модификации, 4, и его эквиваленты на фиг.2-4 каждый работают по меньшей мере частично в области громкости восприятия. Модифицировать звуковой сигнал, 2, работает в области электрических сигналов и формирует модифицированный звуковой сигнал в каждом из примеров фиг.1-4. К тому же в каждом из примеров фиг.1-4, Модифицировать звуковой сигнал, 2, и Сформировать параметры модификации, 4, (или его эквиваленты) модифицируют звуковой сигнал, чтобы уменьшать разницу между его удельной громкостью и целевой удельной громкостью.

В примере с прямой связью фиг.1 последовательность операций или устройство 4 могут включать в себя несколько последовательностей операций и/или устройств: последовательность операций или устройство 6 «Рассчитать целевую удельную громкость», которые рассчитывают целевую удельную громкость в ответ на звуковой сигнал или параметр звукового сигнала, такой как удельная громкость звукового сигнала, последовательность операций или устройство 8 «Рассчитать удельную громкость», которые рассчитывают удельную громкость звукового сигнала в ответ на звуковой сигнал или показатель звуковых сигналов, такой как его накачка, и последовательность операций или устройство 10 «Рассчитать параметры модификации», которые рассчитывают параметры модификации в ответ на удельную громкость и целевую удельную громкость. Рассчитать целевую удельную громкость, 6, может выполнять одну или более функций «F», каждая из которых может иметь параметры функций. Например, он может рассчитывать удельную громкость звукового сигнала, а затем применять одну или более функций F к ней, чтобы предоставить целевую удельную громкость. Это схематически показано на фиг.1 в качестве входного сигнала «Выбрать функцию(и) F и параметр(ы) функции(й)» в последовательность операций или устройство 6. Вместо расчета устройством или последовательностью 6 операций целевая удельная громкость может выдаваться последовательностью операций или устройством хранения (схематически показанными в качестве входного сигнала «Хранимые» в последовательность операций или устройство 10), включенными в или ассоциативно связанными со Сформировать параметры модификации, 4, или источником, внешним по отношению

ко всей последовательности операций, или устройству (схематически показанным в качестве входного сигнала «Внешние» в последовательность операций или устройство 10). Таким образом, параметры модификации основаны, по меньшей мере частично, на расчетах в области (психоакустической) громкости восприятия (то есть по меньшей мере удельной громкости и, в некоторых случаях, расчетах целевой удельной громкости).

Расчеты, выполняемые последовательностями операций или устройствами 6, 8 и 10 (и последовательностями операций или устройствами 12, 14, 10' в примере фиг.2, 6, 14, 10" в примере фиг.3, и 8, 12, 10"" в примере фиг.4), могут выполняться явным и/или неявным образом. Примеры явного выполнения включают в себя (1) справочную таблицу, чьи записи основаны, целиком или частично, на удельной громкости и/или целевой удельной громкости, и/или расчетах параметров модификации, и (2) отражающее ряд решений математическое выражение, которое, по своей природе, основано, целиком или частично, на удельной громкости и/или целевой удельной громкости, и/или параметрах модификации.

Хотя последовательности операций или устройства 6, 8 и 10 расчета по примеру фиг.1 (и последовательности операций или устройства 12, 14, 10' в примере фиг.2, 6, 14, 10" в примере фиг.3, и 8, 12, 10"" в примере фиг.4) схематически показаны и описаны как отдельные, это предназначено только для целей пояснения. Будет понятно, что таковые или все из этих последовательностей операций или устройств могут объединяться в единые последовательности операций или устройстве или по-разному комбинироваться в многочисленных последовательностях операций или устройствах. Например, в компоновке по фиг.9, приведенной ниже, топологии с прямой связью, как в примере по фиг.1, последовательность операций или устройство, которые рассчитывают параметры модификации, делают это в ответ на сглаженную накачку, полученную из звукового сигнала, и целевую удельную громкость. В примере фиг.9 устройство или последовательность операций, которые рассчитывают параметры модификации неявно, рассчитывают удельную громкость звукового сигнала.

В качестве аспекта настоящего изобретения, в примере по фиг.1 и в других примерах вариантов осуществления изобретения, приведенного в материалах настоящей заявки, целевая удельная громкость ($\hat{N}[b, t]$) может рассчитываться масштабированием удельной громкости ($N[b, t]$) одним или более коэффициентами масштабирования. Масштабирование может быть зависящим от времени и частоты масштабным коэффициентом $\Xi[b, t]$, масштабирующим удельную громкость, как в зависимости

$$\hat{N}[b, t] = \Xi[b, t] N[b, t],$$

зависящим от времени, независимым от частоты масштабным коэффициентом $\Phi[t]$, масштабирующим удельную громкость, как в зависимости

$$\hat{N}[b, t] = \Phi[t] N[b, t],$$

независящим от времени, зависящим от частоты масштабным коэффициентом $\Theta[b]$, масштабирующим удельную громкость, как в зависимости

$$\hat{N}[b, t] = \Theta[b] N[b, t], \text{ или}$$

масштабным коэффициентом α , масштабирующим удельную громкость звукового сигнала, как в зависимости

$$\hat{N}[b, t] = \alpha N[b, t],$$

где b - показатель частоты (например, номер полосы), а t - показатель времени

(например, номер кадра). Также могут применяться многочисленные масштабирования, использующие многочисленные экземпляры конкретного масштабирования и/или комбинации конкретных масштабирований. Примеры таких многочисленных масштабирований приведены ниже. В некоторых случаях, как
 5 дополнительно пояснено ниже, масштабирование может быть функцией звукового сигнала или показателей звукового сигнала. В других случаях, как также дополнительно пояснено ниже, когда масштабирование не является функцией показателя звукового сигнала, масштабирование может определяться или подаваться
 10 иным образом. Например, пользователь мог бы выбирать или применять масштабирование с независимым от времени и частоты масштабным коэффициентом α или с независимым от времени, зависящим от частоты масштабным коэффициентом $\Theta[b]$.

Таким образом, целевая удельная громкость может выражаться в качестве одной
 15 или более функций F звукового сигнала или показателя звукового сигнала (удельная громкость является одним из возможных показателей звукового сигнала):

$$\hat{N}[b, t] = F(N[b, t]) .$$

При условии, что функция или функции F являются обратимыми, удельная
 20 громкость ($N[b, t]$) немодифицированного звукового сигнала может рассчитываться в качестве обратной функции или функций F^{-1} целевой удельной громкости ($\hat{N}[b, t]$):

$$N[b, t] = F^{-1}(\hat{N}[b, t]) .$$

Как будет видно ниже, обратная функция или функции F^{-1} рассчитываются в
 25 примерах с обратной связью и смешанной прямой связью/обратной связью по фиг.2 и 4.

Входной сигнал «Выбрать функцию(и) и параметр(ы) функций» для Рассчитать
 30 целевую удельную громкость, 6, показан, чтобы служить признаком, что устройство или последовательность 6 операций может рассчитывать целевую удельную громкость применением одной или более функций в соответствии с одним или более параметров функций. Например, Рассчитать целевую удельную громкость, 8, может рассчитывать функцию или функции « F » удельной громкости звукового сигнала, для того чтобы
 35 определять целевую удельную громкость. Например, входной сигнал «Выбрать функцию(и) и параметр(ы) функции» может выбирать одну или более конкретных функций, которые попадают в один или более вышеприведенных типов масштабирования наряду с одним или более параметров функций, таких как константы (например, масштабные коэффициенты), имеющие отношение к функциям.
 40

Коэффициенты масштабирования, ассоциативно связанные с масштабированием, могут служить в качестве представления целевой удельной громкости ввиду того, что целевая удельная громкость может вычисляться в качестве масштабирования удельной громкости, как указано выше. Таким образом, в примере фиг.9, описанном
 45 ниже и упомянутом выше, справочная таблица может индексироваться коэффициентами масштабирования и накачками из условия, чтобы расчет удельной громкости и целевой удельной громкости был присущ таблице.

С применением справочной таблицы, отражающей ряд решений математического
 50 выражения или некоторой другой технологии, операция Сформировать параметры модификации, 4, (и ее эквивалентные последовательности операций или устройства 4', 4" и 4''' в каждом из примеров фиг.2-4) является такой, что расчеты основаны на области (психоакустической) громкости восприятия, даже если удельная громкость и

целевая удельная громкость могут не рассчитываться явным образом. Есть явная удельная громкость либо есть умозрительная, неявная удельная громкость.

Подобным образом, есть явная целевая удельная громкость либо есть умозрительная, неявная целевая удельная громкость. В любом случае, расчет параметров модификации стремится сформировать параметры модификации, которые модифицируют звуковой сигнал для уменьшения разности между удельной громкостью и целевой удельной громкостью.

В среде воспроизведения, содержащей вторичный мешающий звуковой сигнал, такой как шум, Рассчитать параметры модификации, 10, (и его эквивалентные последовательности операции или устройства 10', 10" и 10''' в каждом из примеров 2-4, соответственно), также может принимать в качестве необязательного входного сигнала показатель такого вторичного мешающего звукового сигнала или сам вторичный мешающий сигнал в качестве одного из своих входных сигналов. Такой необязательный входной сигнал показан на фиг.1 (и на фиг.2-4) пунктирной отвесной линией. Показатель вторичного мешающего сигнала может быть его накачкой, такой как в примере по фиг.17, описанном ниже. Применение показателя мешающего сигнала или самого сигнала (допускается, что мешающий сигнал отдельно доступен для обработки) к последовательности операций или устройствам 10 Рассчитать параметры модификации на фиг.1 (и их эквивалентным последовательностям операций или устройств 10', 10" и 10''' в каждом из примеров фиг.2-4 соответственно) дает возможность надлежащим образом сконфигурировать такие последовательности операций или устройство, чтобы рассчитывать параметры модификации, которые учитывают мешающий сигнал, как дополнительно пояснено ниже под заголовком «Компенсация шума». В примерах по фиг.2-4 расчет удельной громкости частичных тонов предполагает, что подходящий показатель мешающего сигнала подводится не только в соответственный Рассчитать параметры модификации, 10', 10" или 10''', но также и в последовательность операций или устройство 12 «Рассчитать приближение удельной громкости немодифицированного аудио» и или последовательность операций или устройство 14 «Рассчитать приближение целевой удельной громкости», для того чтобы содействовать расчету удельной громкости частичных тонов такой функцией или устройством. В примере с прямой связью фиг.1 удельная громкость частичных тонов не рассчитывается явно - Рассчитать параметры модификации, 10, по фиг.1 рассчитывает надлежащие параметры модификации, чтобы сделать удельную громкость частичных тонов модифицированного аудио приближенной к целевой удельной громкости. Это дополнительно пояснено ниже под заголовком «Компенсация шума», упомянутым выше.

Как упомянуто выше, в каждом из примеров фиг.1-4, параметры М модификации, когда применяются к звуковому сигналу Модификатором 2 звукового сигнала, уменьшают разность между удельной громкостью или удельной громкостью частичных тонов результирующего модифицированного аудио и целевой удельной громкостью. Умозрительно, удельная громкость модифицированного звукового сигнала хорошо приближается к или является такой же, как целевая удельная громкость. Параметры М модификации, например, принимают вид зависящих от времени коэффициентов усиления, применяемых к полосам частот, полученным из гребенки фильтров, или к коэффициентам зависящего от времени фильтра. Соответственно во всех примерах фиг.1-4, Модифицировать звуковой сигнал, 2, например, может быть реализован в качестве множества амплитудных преобразователей масштаба, каждый работает в полосе частот, или зависящего от

времени фильтра (например, многоотводного КИХ-фильтра или многополюсного БИХ-фильтра).

Здесь и где-либо в другом месте в этом документе использование одного и того же номера ссылки указывает, что устройство или последовательность операций могут быть по существу идентичными другому или другим, несущим такой же номер ссылки. Номера ссылок, несущие номера со знаком штриха (например, «10'»), указывают, что устройство или последовательность операций подобны по конструкции или функции, но могут быть модификацией другого или других, несущих такой же базовый номер ссылки или его помеченные знаком штриха варианты.

При определенных ограничениях может быть реализована почти равноценная компоновка с обратной связью по примеру с прямой связью фиг.1. Фиг.2 изображает такой пример, в котором звуковой сигнал также подводится в последовательность операций или устройство 2 Модифицировать звуковой сигнал в сигнальном тракте. Последовательность операций или устройство 2 также принимает параметры М модификации из тракта управления, в котором последовательность операций или устройство 4' Сформировать параметры модификации в компоновке с обратной связью принимает в качестве своего входного сигнала модифицированный звуковой сигнал с выхода Модифицировать звуковой сигнал, 2. Таким образом, в примере фиг.2, скорее модифицированное аудио, нежели немодифицированное аудио, подводится в тракт управления. Последовательность операций или устройство 2 Модифицировать звуковой сигнал и последовательность операций или устройство 4' Сформировать параметры модификации модифицируют звуковой сигнал, чтобы уменьшать разницу между его удельной громкостью и целевой удельной громкостью. Последовательность операций или устройство 4' могут включать в себя несколько функций и/или устройств: последовательность операций или устройство 12 «Рассчитать приближение удельной громкости немодифицированного аудио», последовательность операций или устройство 14 «Рассчитать приближение целевой удельной громкости» и последовательность операций или устройство 10' «Рассчитать параметры модификации», которые рассчитывают параметры модификации.

С ограничением, что функция или функции F обратимы, последовательность операций или устройство 12 оценивает удельную громкость немодифицированного звукового сигнала применением обратной функции F^{-1} к удельной громкости или удельной громкости частичных тонов модифицированного звукового сигнала. Устройство или последовательность 12 операций могут рассчитывать обратную функцию F^{-1} , как описано выше. Это схематически показано на фиг.2 в качестве входного сигнала «Выбрать обратную функцию(и) F^{-1} и параметры функции(й)» в последовательность операций или устройство 12. «Рассчитать приближение целевой удельной громкости» 14 действует, рассчитывая удельную громкость или удельную громкость частичных тонов модифицированного звукового сигнала. Такая удельная громкость или удельная громкость частичных тонов является приближением целевой удельной громкости. Приближение удельной громкости немодифицированного звукового сигнала и приближение целевой удельной громкости используются посредством Рассчитать параметры модификации, 10', для получения параметров М модификации, которые, если применяются к звуковому сигналу посредством Модифицировать звуковой сигнал, 2, уменьшают разницу между удельной громкостью или удельной громкостью частичных тонов модифицированного звукового сигнала и целевой удельной громкостью. Как упомянуто выше, эти параметры М модификации, например, могут принимать вид зависящих от времени

коэффициентов усиления, применяемых к полосам частот гребенки фильтров или коэффициентам зависящего от времени фильтра. В практических вариантах осуществления Рассчитать параметры модификации, 10" цепь обратной связи может привносить задержку между вычислением и применением параметров М модификации.

5 Как упомянуто выше, в среде передачи, содержащей вторичный мешающий звуковой сигнал, такой как шум, каждые из Рассчитать параметры модификации, 10', Рассчитать приближение удельной громкости немодифицированного аудио, 12, и Рассчитать приближение целевой удельной громкости, 14, также могут принимать в
10 качестве необязательного входного сигнала показатель такого вторичного мешающего звукового сигнала или сам вторичный мешающий сигнал в качестве одного из своих входных сигналов, и последовательность операций или устройство 12, и последовательность операций или устройство 14 каждые могут рассчитывать удельную громкость частичных тонов модифицированного звукового сигнала. Такие
15 необязательные входные сигналы показаны на фиг.2 с использованием пунктирных отвесных линий.

Как упомянуто выше, возможны примеры реализаций со смешанной прямой связью/обратной связью аспектов изобретения. Фиг.3 и 4 показывают два примера
20 таких реализаций. В примерах фиг.3 и 4, как и в примерах фиг.1 и 2, звуковой сигнал также подводится в последовательность операций или устройство 2 Модифицировать звуковой сигнал в сигнальном тракте, но Сформировать параметры модификации (4" на фиг.3 и 4'" на фиг.4) в соответственных трактах управления каждые принимают как немодифицированный звуковой сигнал, так и модифицированный звуковой сигнал. В
25 обоих примерах фиг.3 и 4, Модифицировать звуковой сигнал, 2, и Сформировать параметры модификации (4" и 4'", соответственно) модифицируют звуковой сигнал, чтобы уменьшать разницу между его удельной громкостью, которая может быть неявной, и целевой удельной громкостью, которая также может быть неявной.

30 В примере фиг.3 последовательность операций или устройство 4' Сформировать параметры модификации могут включать в себя несколько функций и/или устройств: Рассчитать целевую удельную громкость, 6, как в примере фиг.1, Рассчитать приближение целевой удельной громкости, 14, как в примере с обратной связью фиг.2, и последовательность операций или устройство 10" «Рассчитать параметры
35 модификации». Как в примере фиг.1, в части с прямой связью этого примера со смешанной прямой связью/обратной связью, Рассчитать целевую удельную громкость, 6, может выполнять одну или более функций «F», каждая из которых может иметь параметры функций. Это схематически показано на фиг.3 в качестве
40 входного сигнала «Выбрать функцию(и) F и параметр(ы) функции(й)» в последовательность операций или устройство 6. В части с обратной связью этого примера со смешанной прямой связью/обратной связью модифицированный звуковой сигнал подводится в Рассчитать приближение целевой удельной громкости, 14, как в примере с обратной связью фиг.2. Последовательность операций или устройство 14
45 действует в примере фиг.3, как оно это делает в примере фиг.2, рассчитывая удельную громкость или удельную громкость частичных тонов модифицированного звукового сигнала. Такая удельная громкость или удельная громкость частичных тонов является приближением целевой удельной громкости. Целевая удельная громкость (из
50 последовательности операций или устройства б) и приближение целевой удельной громкости (из последовательности операций или устройства 14) подводятся в Рассчитать параметры модификации, 10", для получения параметров М модификации, которые, если применяются к звуковому сигналу посредством Модифицировать

звуковой сигнал, 2, уменьшают разницу между удельной громкостью немодифицированного звукового сигнала и целевой удельной громкостью. Как упомянуто выше, эти параметры М модификации, например, могут принимать вид зависящих от времени коэффициентов усиления, применяемых к полосам частот гребенки фильтров или коэффициентам зависящего от времени фильтра. В практических вариантах осуществления цепь обратной связи может привносить задержку между вычислением и применением параметров М модификации. Как упомянуто выше, в среде передачи, содержащей вторичный мешающий звуковой сигнал, такой как шум, каждые из Рассчитать параметры модификации, 10", и Рассчитать приближение целевой удельной громкости, 14, также могут принимать в качестве необязательного входного сигнала показатель такого вторичного мешающего звукового сигнала или сам вторичный мешающий сигнал в качестве одного из своих входных сигналов, и последовательность операций или устройство 14 могут рассчитывать удельную громкость частичных тонов модифицированного звукового сигнала. Необязательные входные сигналы показаны на фиг.3 с использованием пунктирных отвесных линий.

Рассчитать параметры модификации, 10", может применять устройство или функцию обнаружения ошибок из условия, чтобы разности между его входными сигналами целевой удельной громкости и приближения целевой удельной громкости настраивали параметры модификации, с тем чтобы уменьшать разности между приближением целевой удельной громкости и «реальной» целевой удельной громкостью. Такие настройки уменьшают разницы между удельной громкостью немодифицированного звукового сигнала и целевой удельной громкостью, которая может быть неявной. Таким образом, параметры М модификации могут адаптироваться на основании ошибки между целевой удельной громкостью, вычисленной в тракте прямой связи из удельной громкости исходного аудио с использованием функции F , и приближением целевой удельной громкости, вычисленным в тракте обратной связи из удельной громкости или удельной громкости частичных тонов модифицированного аудио.

В примере фиг.4 показан альтернативный пример с прямой связью/обратной связью. Этот альтернативный вариант отличается от примера по фиг.3 тем, что в тракте обратной связи предпочтительнее рассчитываются обратная функция или функции F^{-1} , чем функция или функции F , рассчитываемые в тракте прямой связи. В примере фиг.4 последовательность операций или устройство 4' Сформировать параметры модификации могут включать в себя несколько функций и/или устройств: Рассчитать целевую удельную громкость, 8, как в примере с прямой связью фиг.1, Рассчитать приближение целевой удельной громкости немодифицированного аудио, 12, как в примере с обратной связью фиг.2, и «Рассчитать параметры модификации», 10". Рассчитать удельную громкость, 8, как в примере с прямой связью фиг.1, выдает в качестве входного сигнала в Рассчитать параметры модификации, 10", удельную громкость немодифицированного звукового сигнала. Как в примере с обратной связью фиг.2, с ограничением, что функция или функции F обратимы, последовательность операций или устройство 12 оценивает удельную громкость немодифицированного звукового сигнала применением обратной функции F^{-1} к удельной громкости или удельной громкости частичных тонов модифицированного звукового сигнала. Входной сигнал «Выбрать обратную функцию(и) и параметр(ы) обратной функции(й)» для Рассчитать приближение удельной громкости немодифицированного аудио, 12, показан, чтобы служить признаком, что устройство

или последовательность 12 операций могут рассчитывать обратную функцию F^{-1} , как описано выше. Это схематически показано на фиг.4 в качестве входного сигнала «Выбрать обратную функцию(и) F^{-1} и параметр(ы) функции(й)» в последовательность операций или устройство 12. Таким образом, последовательность операций или устройство 12 выдают, в качестве еще одного входного сигнала в Рассчитать параметры модификации, 10'', приближение удельной громкости немодифицированного звукового сигнала.

Как в примерах по фиг.1-3, Рассчитать параметры модификации, 10'', получает параметры М модификации, которые, если применяются к звуковому сигналу посредством Модифицировать звуковой сигнал, 2, уменьшают разность между удельной громкостью немодифицированного звукового сигнала и целевой удельной громкостью, которая является неявной в этом примере. Как упомянуто выше, параметры М модификации, например, могут принимать вид зависящих от времени коэффициентов усиления, применяемых к полосам частот гребенки фильтров или коэффициентам зависящего от времени фильтра. В практических вариантах осуществления цепь обратной связи может привносить задержку между вычислением и применением параметров М модификации. Как упомянуто выше, в среде передачи, содержащей вторичный мешающий звуковой сигнал, такой как шум, каждые из Рассчитать параметры модификации, 10'', и Рассчитать приближение удельной громкости немодифицированного аудио, 12, также могут принимать в качестве необязательного входного сигнала показатель такого вторичного мешающего звукового сигнала или сам вторичный мешающий сигнал в качестве одного из своих входных сигналов, и последовательность операций или устройство 12 могут рассчитывать удельную громкость частичных тонов модифицированного звукового сигнала. Необязательные входные сигналы показаны на фиг.4 с использованием пунктирных отвесных линий.

Рассчитать параметры модификации, 10'', может применять устройство или функцию обнаружения ошибок, из условия чтобы разности между его входными сигналами удельной громкости и приближения удельной громкости формировали выходные сигналы, которые настраивают параметры модификации, с тем чтобы уменьшать разности между приближением удельной громкости и «реальной» удельной громкостью. Так как приближение удельной громкости получается из удельной громкости или удельной громкости частичных тонов модифицированного аудио, которая может рассматриваться как приближение целевой удельной громкости, такие настройки уменьшают разницы между удельной громкостью модифицированного звукового сигнала и целевой удельной громкостью, каковое присуще функции или функциям F^{-1} . Таким образом, параметры М модификации могут адаптироваться на основании ошибки между удельной громкостью, вычисленной в тракте прямой связи из исходного аудио, и приближением удельной громкости, вычисленным с использованием обратной функции или функций F^{-1} , в тракте обратной связи из удельной громкости или удельной громкости частичных тонов модифицированного аудио. Благодаря тракту обратной связи, практические реализации могут привносить задержку между вычислением и применением параметров модификации.

Хотя параметры М модификации в примерах по фиг.1-4, когда подводятся в последовательность операций или устройство 2 Модифицировать Звуковой сигнал, уменьшают разницу между удельной громкостью звукового сигнала и целевой удельной громкостью, в практических вариантах осуществления соответствующие

параметры модификации, формируемые в ответ на один и тот же звуковой сигнал, могут не быть идентичными друг другу.

Хотя некритично и несущественно по отношению к аспектам настоящего изобретения, расчет удельной громкости звукового сигнала или модифицированного звукового сигнала преимущественно может применять технологии, изложенные в упомянутой международной патентной заявке № PCT/US2004/016964, опубликованной как WO 2004/111964 A2, в которой расчет осуществляет выбор из группы из двух или более функций модели удельной громкости, одной или комбинации из двух или более функций модели удельной громкости, выбор которых управляется показателем характеристик входного звукового сигнала. Описание удельной громкости 104 по фиг.1, приведенное ниже, описывает такую компоновку.

В соответствии с дополнительными аспектами изобретения немодифицированный звуковой сигнал и либо (1) параметры модификации, либо (2) целевая удельная громкость или представление целевой удельной громкости (например, масштабные коэффициенты, явно или неявно используемые при расчете целевой удельной громкости), могут сохраняться или передаваться для использования, например, в разделенных временным и/или пространственным образом устройствах или последовательностях операций. Параметры модификации, целевая удельная громкость или представление целевой удельной громкости могут определяться любым подходящим образом, например, как в одном из примеров компоновки с прямой связью, с обратной связью и со смешанной прямой связью/обратной связью по фиг.1-4, как описано выше. На практике компоновка с прямой связью, такая как в примере по фиг.1, является наименее сложной и наиболее быстрой ввиду того, что она избегает расчетов, основанных на модифицированном звуковом сигнале. Пример передачи или сохранения немодифицированного аудио и параметров модификации показан на фиг.5 наряду с тем, что пример передачи или сохранения немодифицированного аудио и целевой удельной громкости или представления целевой удельной громкости показан на фиг.6.

Компоновка, такая как в примере по фиг.5, может использоваться для временного и/или пространственного отделения применения параметров модификации к звуковому сигналу от формирования таких параметров модификации. Компоновка, такая как в примере по фиг.6, может использоваться для временного и/или пространственного отделения обоих, формирования и применения параметров модификации, от формирования целевой удельной громкости или ее представления. Оба типа компоновок делают возможными простые недорогие компоновки воспроизведения или приема, которые избегают сложности формирования параметров модификации или формирования целевой удельной громкости. Хотя компоновка типа по фиг.5 проще, чем компоновка типа по фиг.6, компоновка фиг.6 имеет преимущество, что информация, требуемая для сохранения или передачи, может быть гораздо меньшей, в частности, когда представление целевой удельной громкости, такое как один или более масштабных коэффициентов, сохраняется или передается. Такое сокращение хранения или передачи информации, в частности, может быть полезным в звуковых средах с низкой скоростью передачи битов.

Соответственно дополнительными аспектами настоящего изобретения является предоставление устройства или последовательности операций, которые (1) принимают или воспроизводят, из устройства или последовательности операций сохранения или передачи, параметры М модификации и применяют их к звуковому сигналу, который также принимается, или (2) которые принимают или воспроизводят, из устройства или

последовательности операций сохранения или передачи, целевую удельную громкость или представление целевой удельной громкости, формируют параметры М модификации, применяя целевую удельную громкость или ее представление к звуковому сигналу, который также принимается (или к показателю звукового сигнала, 5 такому как его удельная громкость, который может получаться из звукового сигнала), и применяют параметры М модификации к принятому звуковому сигналу. Такие устройства или последовательности операций могут характеризоваться как последовательности операций декодирования или декодеры; наряду с тем, что 10 устройства или последовательности операций, требуемые для формирования сохраняемой или передаваемой информации, могут характеризоваться как последовательности операций кодирования или кодировщики. Такие последовательности операций кодирования или кодировщики являются теми частями примеров компоновок фиг.1-4, которые являются используемыми для формирования 15 информации, требуемой соответственными последовательностями операций декодирования или декодерами. Такие процессоры декодирования или декодеры могут быть ассоциативно связанными или работающими практически с любым типом последовательности операций или устройства, которые обрабатывают и/или 20 воспроизводят звук.

В одном из аспектов изобретения, как в примере по фиг.5, немодифицированный звуковой сигнал и параметры М модификации, например, сформированные последовательностью операций формирования или генератором параметров модификации, таким как Сформировать параметры модификации, 4, по фиг.1, 4' по 25 фиг.2, 4" по фиг.3 или 4''' по фиг.4 могут подводиться в любое подходящее устройство или функцию 16 сохранения или передачи («Сохранить или Передать»). В случае использования примера с прямой связью по фиг.1 в качестве последовательности операций кодирования или кодировщика, Модифицировать звуковой сигнал, 2, не 30 потребовался бы для формирования модифицированного аудио и мог бы быть опущен, если нет потребности предоставлять модифицированное аудио во временном или пространственном местоположении кодировщика или последовательности операций кодирования. Сохранить или передать, 16, например, может включать в себя 35 любые подходящие оптические или твердотельные устройства хранения и воспроизведения или любые подходящие проводные или беспроводные устройства передачи и приема, их выбор не является критичным по отношению к изобретению. Воспроизведенные или принятые параметры модификации затем могут подводиться в Модифицировать звуковой сигнал, 2, типа, примененного в примерах по фиг.1-4, для 40 того чтобы модифицировать воспроизводимый или принимаемый звуковой сигнал, так что его удельная громкость приближается к целевой удельной громкости, или присуща компоновке, в которой получались параметры модификации. Параметры модификации могут сохраняться или передаваться любым из различных способов. Например, они могут сохраняться или передаваться в качестве метаданных, 45 сопровождающих звуковой сигнал, они могут отправляться в отдельных трактах или каналах, они могут стенографически кодироваться в аудио, они могут мультиплексироваться и т.п. Использование параметров модификации для модификации звукового сигнала может быть необязательным и, если необязательно, их использование может быть выбираемым, например, пользователем. Например, 50 параметры модификации, если применяются к звуковому сигналу, могли бы уменьшать динамический диапазон звукового сигнала. Следует ли применять или не применять такое уменьшение динамического диапазона могло бы быть выбираемым

пользователем.

В еще одном аспекте изобретения, как в примере по фиг.6, немодифицированный звуковой сигнал и целевая удельная громкость или представление целевой удельной громкости могут подводиться в любые подходящие устройство или функцию 16 сохранения или передачи («Сохранить или передать»). В случае использования конфигурации с прямой связью, такой как пример по фиг.1, в качестве последовательности операций кодирования или кодировщика, ни последовательность операций или устройство типа Рассчитать параметры модификации, 10, ни последовательность операций или устройство типа Модифицировать звуковой сигнал, 2, не потребовались бы и могли бы быть опущены, если нет потребности предоставлять параметры модификации или модифицированное аудио во временном или пространственном местоположении кодировщика или последовательности операций кодирования. Как в случае примера фиг.5, Сохранить или передать, 16, например, может включать в себя любые подходящие оптические или твердотельные устройства хранения и воспроизведения, или любые подходящие проводные или беспроводные устройства передачи и приема, их выбор не является критичным по отношению к изобретению. Воспроизведенные или принятые целевая удельная громкость или представление целевой удельной громкости затем могут подводиться, наряду с немодифицированным аудио, в Рассчитать параметры модификации, 10, типа, примененного в примерах по фиг.3, для того чтобы предоставить параметры М модификации, которые затем могут подводиться в Модифицировать звуковой сигнал, 2, типа, применяемого в примерах по фиг.1-4, для того чтобы модифицировать воспроизводимый или принимаемый сигнал, так что его удельная громкость приближается к целевой удельной громкости или присуща компоновке, в которой получались параметры модификации. Хотя целевая удельная громкость или ее представление могут наиболее легко получаться в последовательности операций кодирования или кодировщике примерного типа фиг.10, целевая удельная громкость или ее представление либо приближение к целевой удельной громкости или его представление могут получаться в последовательности операций кодирования или кодировщике примерных типов фиг. со 2 по 4 (приближения рассчитываются в последовательностях операций или устройствах 14 по фиг.2 и 3 и в последовательности операций или устройстве 12 по фиг.4). Целевая удельная громкость или ее представление могут сохраняться или передаваться любым из различных способов. Например, она может сохраняться или передаваться в качестве метаданных, сопровождающих звуковой сигнал, она может отправляться в отдельных трактах или каналах, она может стенографически кодироваться в аудио, она может мультиплексироваться и т.п. Использование параметров модификации, полученных из хранимых или переданных целевой удельной громкости, или представления для модификации звукового сигнала может быть необязательным и, если необязательно, их использование может быть выбираемым, например, пользователем. Например, параметры модификации, если применяются к звуковому сигналу, могли бы уменьшать динамический диапазон звукового сигнала. Следует ли применять или не применять такое уменьшение динамического диапазона могло бы быть выбираемым пользователем.

При реализации раскрытого изобретения в качестве цифровой системы наиболее практичной является конфигурация с прямой связью, а потому примеры таких конфигураций подробно описаны ниже, подразумевается, что объем изобретения не ограничен таким образом.

На всем протяжении этого документа термины, такие как «фильтр» или «гребенка фильтров», используются в материалах настоящей заявки для включения в состав по существу любой разновидности рекурсивной и нерекурсивной фильтрации, такой как БИХ-фильтры или преобразования, и «фильтрованная» информация является

результатом применения таких фильтров. Варианты осуществления, описанные ниже, применяют гребенки фильтров, реализованные преобразованиями.

Фиг.7 изображает большие подробности примерного варианта осуществления аспекта изобретения, воплощенного в компоновке с прямой связью. Аудио сначала проходит через функцию или устройство 100 гребенки фильтров анализа («Гребенка фильтров анализа»), которые разделяют звуковой сигнал на множество полос частот (отсюда фиг.5 показывает многочисленные выходные сигналы из Гребенки фильтров анализа, 100, каждый выходной сигнал представляет полосу частот, которую выходной сигнал проносит через различные функции или устройства вплоть до гребенки фильтров синтеза, которая суммирует полосы в объединенный широкополосный сигнал, как дополнительно описано ниже). Характеристика фильтра, ассоциативно связанного с каждой полосой частот в Гребенке фильтров анализа, 100, предназначена для имитации характеристики в конкретном местоположении базилярной мембраны во внутреннем ухе. Выходной сигнал каждого фильтра в Гребенке фильтров анализа, 100, затем проходит в пропускающий фильтр или функцию 101 пропускающего фильтра («Пропускающий фильтр»), которые имитируют эффект фильтрации передачи аудио через наружное и среднее ухо. Если должна была измеряться только громкость аудио, пропускающий фильтр мог бы применяться до гребенки фильтров анализа, но, так как выходные сигналы гребенки фильтров анализа используются для синтеза модифицированного аудио, полезно применять пропускающий фильтр после гребенки фильтров. Выходные сигналы Пропускающего фильтра, 101, затем проходят в функцию или устройство 102 накачки («Накачка»), выходные сигналы которых имитируют распределение энергии по базилярной мембране. Значения энергии накачки могут сглаживаться по времени функцией или устройством 103 сглаживания («Сглаживание»). Постоянные времени функции сглаживания заданы в соответствии с потребностями желательного применения. Сглаженные сигналы накачки по существу преобразуются в удельную громкость в функции или устройстве 104 удельной громкости («Удельная громкость (SL)»). Удельная громкость представлена в единицах сонов на единичную частоту. Составляющая удельной громкости, ассоциативно связанная с каждой полосой, проходит в функцию или устройство 105 модификации удельной громкости («Модификация SL»). Модификация 105 SL принимает в качестве своего входного сигнала исходную удельную громкость, а затем выводит требуемую или «целевую» удельную громкость, которая согласно аспекту настоящего изобретения предпочтительно является функцией исходной удельной громкости (смотрите следующий ниже заголовок, озаглавленный «Целевая удельная громкость»).

Модификация SL, 105 может работать независимо в каждой полосе, или может существовать взаимная зависимость между или среди полос (частотное сглаживание, как подсказано линиями перекрестных соединений на фиг.7), зависящая от требуемого результата. Принимая в качестве своих входных сигналов сглаженные составляющие полос частот накачки из Накачки, 102, и целевую удельную громкость из Модификации SL, 105, функция или устройство 106 решателя коэффициентов усиления («Решатель коэффициентов усиления») определяет коэффициент усиления, которому необходимо применяться к каждой полосе выходного сигнала Гребенки фильтров

анализа, 100, для того чтобы преобразовывать измеренную удельную громкость в целевую удельную громкость. Решатель коэффициентов усиления может быть реализован различными способами. Например, Решатель коэффициентов усиления может включать в себя итерационную последовательность операций, такую как в

5 способе, который раскрыт в упомянутой международной патентной заявке № РСТ/US2004/016964, опубликованной как WO 2004/111964 A2, или в качестве альтернативы справочную таблицу. Хотя коэффициенты на полосу, формируемые Решателем коэффициентов усиления, 106, могут дополнительно сглаживаться по

10 времени необязательными функцией или устройством 107 сглаживания («Сглаживание»), для того чтобы минимизировать артефакты восприятия, предпочтительно, чтобы временное сглаживание применялось где-то в другом месте во всей последовательности операций или устройстве, как описано где-либо в другом месте. В заключение коэффициенты усиления применяются к соответственным

15 полосам Гребенки фильтров анализа, 100 через соответственную функцию мультипликативного объединения или объединитель 108, и обработанное или «модифицированное» аудио синтезируется из модифицированных коэффициентами усиления полос в функции или устройстве 110 гребенки фильтров синтеза («Гребенка фильтров синтеза»). В дополнение выходные сигналы из гребенки фильтров анализа могут задерживаться функцией или устройством 109 задержки («Задержка») до

20 применения коэффициентов усиления, для того чтобы компенсировать любое запаздывание, ассоциативно связанное с вычислением коэффициентов усиления. В качестве альтернативы вместо расчета коэффициентов усиления для использования при применении модификаций коэффициентов усиления в полосах частот, Решатели коэффициентов усиления, 106, могут рассчитывать коэффициенты фильтра, которые управляют зависящим от времени фильтром, таким как многоотводный КИХ-фильтр или многополюсный БИХ-фильтр. Для простоты в описании аспекты изобретения,

30 главным образом, описаны в качестве использования коэффициентов усиления, применяемых к полосам частот, подразумевается, что коэффициенты фильтра и зависящие от времени фильтры также могут применяться в практических вариантах осуществления.

В конкретных вариантах осуществления обработка аудио может выполняться в

35 цифровой области. Соответственно звуковой входной сигнал обозначен дискретной временной последовательностью $x[n]$, которая была дискретизирована из источника аудио на некоторой частоте f_s выборки. Предполагается, что последовательность $x[n]$ была надлежащим образом масштабирована, так что среднеквадратическая мощность $x[n]$ в децибелах, заданная посредством

40

$$RMS_{dB} = 10 \log_{10} \left(\frac{1}{L} \sum_{n=0}^L x^2[n] \right)$$

равна уровню звукового давления в дБ, на котором аудио прослушивается

45 человеком-слушателем. В дополнение звуковой сигнал предполагается монофоническим для простоты описания.

Гребенка фильтров анализа, 100, Пропускающий фильтр, 101, Накачка, 102, Удельная громкость, 104, Модификация удельной громкости, 105, Решатель коэффициентов усиления, 106, и Гребенка фильтров синтеза, 110, могут быть описаны

50 более подробно, как изложено ниже.

Гребенка фильтров, 100

Звуковой входной сигнал подводится в гребенку фильтров или функцию 100

гребенки фильтров анализа («Гребенка фильтров анализа»). Каждый фильтр в Гребенке фильтров анализа, 100, предназначен для имитации частотной характеристики в конкретном местоположении вдоль базилярной мембраны во внутреннем ухе. Гребенка 100 фильтров может включать в себя набор линейных фильтров, чьи полосы пропускания и разнесение постоянны по шкале эквивалентной прямоугольной полосы пропускания (ERB), как определено Муром, Глазбергом и Баером (B. C. J. Moore, B. Glasberg, T. Baer, «A Model for the Prediction of Thresholds, Loudness, and Partial Loudness», приведенном выше).

Хотя шкала частот ERB ближе соответствует человеческому восприятию и показывает улучшенное качество функционирования при создании измерений объективной громкости, которые соответствуют результатам субъективной громкости, шкала частот Барка может применяться с пониженным качеством функционирования.

Для центральной частоты f в герцах ширина одной полосы ERB в герцах может быть приближенно выражена как:

$$ERB(f) = 24.7(4.37f/1000 + 1) \quad (1)$$

По этой зависимости определяется криволинейная шкала частот из условия, чтобы любая точка по криволинейной шкале, соответствующая ERB в единицах криволинейной шкалы, была равна единице. Функция для преобразования из линейной частоты в Герцах в эту шкалу частот ERB получается интегрированием аналога уравнения 1:

$$HzToERB(f) = \int \frac{1}{24.7(4.37f/1000 + 1)} df = 21.4 \log_{10}(4.37f/1000 + 1) \quad (2a)$$

Также полезно выражать преобразование из шкалы ERB обратно в линейную шкалу частот посредством решения уравнения 2a в отношении f :

$$ERBToHz(e) = f = \frac{1000}{4.37} 10^{(e/21.4-1)}, \quad (2b)$$

где e имеет место в единицах шкалы ERB. Фиг.9 показывает зависимость между шкалой ERB и частотой в Герцах.

Гребенка фильтров анализа, 100, может включать в себя B слуховых фильтров, указываемых ссылкой как полосы, на центральных частотах $f_c[1] \dots f_c[B]$, равномерно разнесенных вдоль шкалы ERB. Более точно,

$$f_c[1] = f_{\min} \quad (3a)$$

$$f_c[b] = f_c[b-1] + ERBToHz(HzToERB(f_c[b-1]) + \Delta) \quad b = 2 \dots B \quad (3b)$$

$$f_c[B] < f_{\max} \quad (3c)$$

где Δ - требуемое разнесение ERB Гребенки фильтров анализа, 100 и где $f_{\min}$ и $f_{\max}$ - требуемые минимальные и максимальные центральные частоты соответственно.

Можно выбрать $\Delta=1$, и, учитывая частотный диапазон, на котором чувствительно человеческое ухо, можно установить $f_{\min}=50$ Гц и $f_{\max}=20000$ Гц. С такими параметрами, например, применение уравнений 3a-с дает $B=40$ слуховых фильтров.

Амплитудно-частотная характеристика каждого слухового фильтра может характеризоваться сферической экспоненциальной функцией, как предложено Муром и Глазбергом. Более точно, амплитудная характеристика фильтра с центральной частотой $f_c[b]$ может вычисляться в качестве:

$$H_b(f) = (1 + pg)e^{-pg} \quad (4a)$$

где

$$g = \left| \frac{f - f_c[b]}{f_c[b]} \right|, \quad (4b)$$

$$p = \frac{4f_c[b]}{ERB(f_c[b])} \quad (4c)$$

Амплитудные характеристики таких B слуховых фильтров, которые приближенно равняются определению критических полос по шкале ERB, показаны на фиг.10.

Операции фильтрации Гребенки фильтров анализа, 100, могут удовлетворительно приближенно выражаться с использованием дискретного преобразования Фурье конечной длины, обычно указываемого ссылкой как кратковременное дискретное преобразование Фурье (STDFT), так как реализация, выполняющая фильтры на частоте выборки звукового сигнала, указываемая ссылкой как полночастотная реализация, предполагается дающей большее временное разрешение, чем необходимо для точных измерений громкости. Посредством использования STDFT вместо полночастотной реализации может достигаться улучшение эффективности и снижение вычислительной сложности.

STDFT входного звукового сигнала $x[n]$ определено в качестве:

$$X[k, t] = \sum_{n=0}^{N-1} w[n] x[n + tT] e^{-j \frac{2\pi k n}{N}}, \quad (5a)$$

где k - индекс частоты, t - индекс временного интервала, N - размер ДПФ (дискретного преобразования Фурье, DFT), T - размер скачка, а $w[n]$ - длина окна N , нормализованного так, что

$$\sum_{n=0}^{N-1} w^2[n] = 1 \quad (5b)$$

Отметим, что переменная t в уравнении 5a является дискретным индексом, представляющим временной интервал STDFT, в противоположность измерению времени в секундах. Каждое приращение в t представляет скачок на T отсчетов вдоль сигнала $x[n]$. Последующие ссылки на индекс t предполагают это определение. Несмотря на то что разные настройки параметров и формы окна могут использоваться в зависимости от деталей реализаций, для $f_s=44100$ Гц, выбор $N=2048$, $T=1024$, и вынуждение $w[n]$ быть окном Хеннинга обеспечивает достаточный баланс временного и частотного разрешения. STDFT, описанное выше, может быть более эффективным с использованием быстрого преобразования Фурье (БПФ, FFT).

Вместо STDFT для реализации гребенки фильтров анализа может использоваться модифицированное дискретное косинусное преобразование (MDCT). MDCT-преобразование, обычно используемое в кодировщиках связанного с восприятием аудио, таких как AC-3 системы Долби. Если раскрытая система реализуется с помощью такого перцепционно кодированного аудио, раскрытые измерение и модификация громкости могут более эффективно реализовываться обработкой существующих коэффициентов MDCT кодированного аудио, тем самым устраняя необходимость выполнять преобразование гребенки фильтра анализа. MDCT входного звукового сигнала $x[n]$ задано посредством:

$$X[k, t] = \sum_{n=0}^{N-1} w[n] x[n + tT] \cos((2\pi / N)(k + 1/2)(n + n_0)),$$

$$\text{где } n_0 = \frac{(N/2) + 1}{2} \quad (6)$$

Обычно размер T скачка выбирается, чтобы быть точно половиной длины N преобразования, так что возможна безукоризненная реконструкция сигнала $x[n]$.

Пропускающий фильтр, 101

Выходные сигналы Гребенки фильтров анализа, 100, подводятся в пропускающий фильтр или функцию 101 пропускающего фильтра («Пропускающий фильтр»), которые фильтруют каждую полосу гребенки фильтров в соответствии с передачей аудио через наружное и среднее ухо. Фиг.8 изображает одну из пригодных амплитудно-частотных характеристик пропускающего фильтра, $P(f)$, на ширине диапазона слышимых частот. Характеристика является единицей ниже 1 кГц, а выше 1 кГц, следуют инверсии порога слышимости, который задан в стандарте ISO226, с пороговым значением, нормализованным для равенства единице на 1 кГц.

Накачка, 102

Для того чтобы вычислять громкость входного звукового сигнала, необходим показатель кратковременной энергии звуковых сигналов в каждом фильтре Гребенки фильтров анализа, 100, после применения Пропускающего фильтра, 101. Этот зависящий от времени и частоты показатель указывается ссылкой как накачка. Выходной сигнал кратковременной энергии каждого фильтра в Гребенке фильтров анализа, 100, может приближенно выражаться в Функции накачки, 102, посредством перемножения характеристик фильтров в частотной области со спектром мощности входного сигнала:

$$E[b, t] = \frac{1}{N} \sum_{k=0}^{N-1} |H_b[k]|^2 |P[k]|^2 |X[k, t]|^2, \quad (7)$$

где b - номер полосы, t - номер кадра, а $H_b[k]$ и $P[k]$ - частотные характеристики слухового фильтра и пропускающего фильтра, соответственно дискретизированные на частоте, соответствующей индексу k приемника STDFT или MDCT. Должно быть отмечено, что разновидности для амплитудной характеристики слуховых фильтров, иные чем заданные в уравнениях 4а-с, могут использоваться в уравнении 7 для достижения подобных результатов. Например, упомянутая международная заявка № PCT/US2004/016964, опубликованная как WO 2004/111964 A2, описывает два альтернативных варианта: слуховой фильтр, характеризуемый функцией БИХ-преобразования 12^{ого} порядка, и приближение недорогого полосового фильтра «с крутым срезом».

Итак, выходной сигнал Функции накачки, 102, является представлением в частотной области энергии E в соответственных полосах b ERB за период t времени.

Усреднение по времени («Сглаживание»), 103

Для некоторых применений раскрытого изобретения, как описано ниже, может быть желательным сглаживать накачку $E[b, t]$ до ее преобразования в удельную громкость. Например, сглаживание может выполняться рекурсивно в функции 103 сглаживания согласно уравнению:

$$\bar{E}[b, t] = \lambda_b \bar{E}[b, t] + (1 - \lambda_b) E[b, t], \quad (8)$$

где постоянные времени λ_b в каждой полосе b выбираются в соответствии с

требуемым применением. В большинстве случаев постоянные времени преимущественно могут выбираться, чтобы быть пропорциональными времени интегрирования человеческого восприятия громкости в пределах полосы b . Уотсон и Гендель выполняли эксперименты, демонстрирующие, что это время интегрирования находится в пределах диапазона в 150-175 мс на низких частотах (125-200 Гц) и 40-60 мс на высоких частотах (Charles S. Watson and Roy W. Gengel, «Signal Duration and Signal Frequency in Relation to Auditory Sensitivity» *Journal of the Acoustical Society of America*, Vol. 46, No. 4 (Part 2), 1969, pp. 989-997 (Чарли С., Уотсон и Рой В. Гендель, «Длительность сигнала и частота сигнала относительно слуховой чувствительности» *Журнал Акустического сообщества Америки*, том 46, № 4 (часть 2), 1969 год, стр. 989-997)).

Удельная громкость, 104

В преобразователе или функции 104 преобразования удельной громкости («Удельная громкость») каждая полоса накачки преобразуется в значение составляющей удельной громкости, которое измеряется в сонах на ERB.

Вначале при вычислении удельной громкости уровень накачки в каждой полосе $\bar{E}[b, t]$ может преобразовываться в эквивалентный уровень накачки на 1 кГц, как задано кривыми равной громкости по ISO 226 (фиг.11), нормализованным пропускающим фильтром $P(z)$ фиг.12):

$$\bar{E}_{1kHz}[b, t] = T_{1kHz}(\bar{E}[b, t], f_c[b]), \quad (9)$$

где $T_{1kHz}(E, f)$ - функция, которая формирует уровень на 1 кГц, который равен по громкости уровню E на частоте f . На практике, $T_{1kHz}(E, f)$ реализуется в качестве интерполяции справочной таблицы кривых равной громкости, нормализованных пропускающим фильтром. Преобразование эквивалентных уровней на 1 кГц упрощает последующий расчет удельной громкости.

Затем удельная громкость в каждой полосе может быть вычислена в качестве:

$$N[b, t] = \alpha[b, t]N_{NB}[b, t] + (1 - \alpha[b, t])N_{WB}[b, t], \quad (10)$$

где $N_{NB}[b, t]$ и $N_{WB}[b, t]$ - значения удельной громкости на основании узкополосной или широкополосной модели прохождения сигнала соответственно. Значение $\alpha[b, t]$ - является коэффициентом интерполяции, лежащим между 0 и 1, который вычисляется из звукового сигнала. Упомянутая международная заявка № PCT/US2004/016964, опубликованная как WO 2004/111964 A2, описывает технологию для расчета $\alpha[b, t]$ по спектральной неравномерности накачки. Она также более подробно описывает «узкополосные» и «широкополосные» модели прохождения сигнала.

Значения $N_{NB}[b, t]$ и $N_{WB}[b, t]$ узкополосной и широкополосной удельной громкости могут оцениваться по преобразованной накачке с использованием экспоненциальных функций:

$$N_{NB}[b, t] = \begin{cases} G_{NB} \left(\left(\frac{\bar{E}_{1kHz}[b, t]}{TQ_{1kHz}} \right)^{p_{NB}} - 1 \right), & \bar{E}_{1kHz}[b, t] > TQ_{1kHz} \\ 0, & \text{иначе} \end{cases} \quad (11a)$$

$$N_{WB}[m, t] = \begin{cases} G_{WB} \left(\left(\frac{\bar{E}_{1kHz}[b, t]}{TQ_{1kHz}} \right)^{p_{WB}} - 1 \right), & \bar{E}_{1kHz}[b, t] > TQ_{1kHz} \\ 0, & \text{иначе} \end{cases}, \quad (11b)$$

где TQ_{1kHz} - уровень накачки при пороговом значении в тишине для тона в 1 кГц.

По кривым равной громкости (фиг.11 и 12) $TQ_{1\text{kHz}}$ равен 4,2 дБ. Отмечаем, что обе из этих функций удельной громкости равны нулю, когда накачка равна пороговому значению в тишине. Для накачек, больших чем пороговое значение в тишине, обе функции монотонно возрастают со степенной зависимостью в соответствии с законом Стивенса об ощущении интенсивности. Показатель степени для узкополосной функции выбирается, чтобы быть большим, чем таковой у широкополосной функции, заставляя узкополосную функцию возрастать быстрее, чем широкополосная функция. Отдельный набор показателей степени β и коэффициентов усиления G для узкополосного и широкополосного случаев выбирается, чтобы соответствовать экспериментальным данным о росте громкости для тонов и шума.

Мур и Глазберг выдвинули в качестве предположения, что удельная громкость должна быть равной некоторому небольшому значению вместо нуля, когда накачка находится на пороге слышимости. Удельная громкость затем должна монотонно уменьшаться до нуля, по мере того как накачка уменьшается до нуля. Обоснование состоит в том, что порог слышимости является вероятностным порогом (точкой, в которой тон обнаруживается 50% времени) и что количество тонов, каждый при пороговом значении, представленных совместно, может суммироваться в звук, который слышим в большей степени, чем любой из отдельных тонов. В раскрытой заявке пополнение функций удельной громкости этим свойством имеет дополнительное преимущество побуждения решателя коэффициентов усиления, обсужденного ниже, вести себя более уместно, когда накачка находится возле порогового значения. Если удельная громкость определена равной нулю, когда накачка находится на или ниже порогового значения, то уникального решения для решателя коэффициентов усиления не существует для накачек на или ниже порогового значения. Если, с другой стороны, удельная громкость определена монотонно повышающейся для всех значений накачки, больших чем или равных нулю, как предложено Муром и Глазбергом, то уникальное решение существует. Масштабирование громкости, большее чем единица, всегда будет иметь следствием коэффициент усиления, больший чем единица, и наоборот. Функции удельной громкости в уравнениях 11a и 11b могут быть видоизменены, чтобы иметь требуемое свойство, согласно:

$$N_{NB}[b, t] = \begin{cases} G_{NB} \left(\left(\frac{\bar{E}_{1kHz}[b, t]}{TQ_{1kHz}} \right)^{\eta_{NB}} - 1 \right), & \bar{E}_{1kHz}[b, t] > \lambda TQ_{1kHz} \\ \exp \left\{ K_{NB} \left(-\log \left(\frac{\bar{E}_{1kHz}[b, t]}{TQ_{1kHz}} \right) + C_{NB} \right)^{\eta_{NB}} \right\}, & \text{иначе} \end{cases} \quad (11c)$$

$$N_{WB}[m, t] = \begin{cases} G_{WB} \left(\left(\frac{\bar{E}_{1kHz}[b, t]}{TQ_{1kHz}} \right)^{\eta_{WB}} - 1 \right), & \bar{E}_{1kHz}[b, t] > \lambda TQ_{1kHz} \\ \exp \left\{ K_{WB} \left(-\log \left(\frac{\bar{E}_{1kHz}[b, t]}{TQ_{1kHz}} \right) + C_{WB} \right)^{\eta_{WB}} \right\}, & \text{иначе} \end{cases} \quad (11d)$$

где константа λ является большей, чем единица, показатель η степени является меньшим, чем единица, а константы K и C выбираются так, что функция удельной громкости и ее первая производная являются непрерывными в точке $\bar{E}_{1kHz}[b, t] = \lambda TQ_{1kHz}$.

Из удельной громкости, общая или «полная» громкость $L[t]$ задается суммой удельной громкости по всем полосам b :

$$L[t] = \sum_b N[b, t] \quad (12)$$

Модификация удельной громкости, 105

В функции 105 модификации удельной громкости («Модификация удельной громкости») целевая удельная громкость, указываемая ссылкой как $\hat{N}[b, t]$, может рассчитываться по удельной громкости из SL 104 (фиг.7) различными способами, зависящими от требуемого применения всего устройства или последовательности операций. Как более подробно описано ниже, целевая удельная громкость может рассчитываться с использованием масштабного коэффициента α , например, в случае регулировки уровня громкости. Смотрите уравнение 16, приведенное ниже, и его ассоциативно связанное описание. В случае автоматической регулировки усиления (APU) и регулировки динамического диапазона (DRC) целевая удельная громкость может рассчитываться с использованием отношения требуемой выходной громкости ко входной громкости. Смотрите уравнения 17 и 18, приведенные ниже, и их ассоциативно связанные описания. В случае динамической коррекции целевая удельная громкость может рассчитываться с использованием зависимости, изложенной в уравнении 23, и ее ассоциативно связанном описании.

Решатель коэффициентов усиления, 106

В этом примере для каждой полосы b и каждого интервала t времени, Решатель коэффициентов усиления, 106, принимает в качестве своих входных сигналов сглаженную накачку $\bar{E}[b, t]$ и целевую удельную громкость $\hat{N}[b, t]$ и формирует коэффициенты $G[b, t]$ усиления, используемые впоследствии для модификации аудио. С допущением, что функция $\Psi\{\cdot\}$ представляет нелинейное преобразование из накачки в удельную громкость, из условия чтобы

$$\hat{N}[b, t] = \Psi\{\bar{E}[b, t]\}, \quad (13)$$

Решатель коэффициентов усиления находит $G[b, t]$, так что

$$\hat{N}[b, t] = \Psi\{G^2[b, t]\bar{E}[b, t]\}. \quad (14a)$$

Решатели коэффициентов усиления, 106, определяют зависящие от частоты и времени коэффициенты усиления, которые, когда применяются к исходной накачке, дают в результате громкость, которая в соответствии с идеалом равна требуемой целевой удельной громкости. На практике, Решатели коэффициентов усиления определяют зависящие от частоты и времени коэффициенты усиления, которые, когда применяются к варианту в частотной области звукового сигнала, дают в результате модификацию звукового сигнала, для того чтобы уменьшать разность между его удельной громкостью и целевой удельной громкостью. Умозрительно модификация является такой, что модифицированный звуковой сигнал имеет удельную громкость, которая является хорошим приближением целевой удельной громкости. Решение для уравнения 14a может быть реализовано многообразием способов. Например, если существует отражающее ряд решений математическое выражение для инверсии удельной громкости, представленной посредством $\Psi^{-1}\{\cdot\}$, то коэффициенты усиления могут вычисляться непосредственно перекомпоновкой уравнения 14a:

$$G[b, t] = \sqrt{\frac{\Psi^{-1}(\hat{N}[b, t])}{\bar{E}[b, t]}} \quad (14b)$$

В качестве альтернативы, если отражающего ряд решений математического выражения для $\Psi^{-1}\{\cdot\}$ не существует, может использоваться итерационный подход, в котором в течение каждой итерации уравнение 14a оценивается с использованием текущей оценки коэффициентов усиления. Результирующая удельная громкость сравнивается с требуемой целевой, и коэффициенты усиления обновляются на основании ошибки. Если коэффициенты усиления обновляются надлежащим образом,

они будут сходиться к требуемому решению. Другой способ включает в себе предварительное вычисление функции $\Psi^{-1}\{\cdot\}$ для диапазона значений накачки в каждой полосе, чтобы создавать справочную таблицу. Из этой справочной таблицы получают приближение обратной функции $\Psi^{-1}\{\cdot\}$, а коэффициенты усиления, в таком случае, могут вычисляться из уравнения 14b. Как упомянуто ранее, целевая удельная громкость может быть представлена масштабированием удельной громкости;

$$\hat{N}[b, l] = \Xi[b, l] N[b, l] \quad (14c)$$

Подстановка уравнения 13 в 14с, а затем, 14с в 14b дает альтернативное выражение для коэффициентов усиления:

$$G[b, l] = \sqrt{\frac{\Psi^{-1}(\Xi[b, l] \Psi(\bar{E}[b, l]))}{\bar{E}[b, l]}} \quad (14d)$$

Мы видим, что коэффициенты усиления могут безукоризненно выражаться в качестве функции накачки $\bar{E}[b, l]$ и масштабирования $\Xi[b, l]$ удельной громкости. Поэтому, коэффициенты могут вычисляться посредством оценки по 14d или эквивалентной справочной таблицы без явного вычисления когда-либо удельной громкости или целевой удельной громкости в качестве промежуточных значений. Однако, эти значения вычисляются неявно благодаря использованию уравнения 14d. Могут быть изобретены другие эквивалентные способы для вычисления параметров модификации через явное или неявное вычисление удельной громкости и целевой удельной громкости, и это изобретение подразумевается покрывающим все такие способы.

Гребенка фильтров синтеза, 110

Как описано выше, Гребенка фильтров анализа, 100, может быть эффективно реализована благодаря использованию кратковременного дискретного преобразования Фурье (STDFT) или модифицированного дискретного косинусного преобразования, а STDFT или MDCT могут использоваться подобным образом для реализации Гребенки фильтров синтеза, 110. Более точно, с допущением, что $X[k, l]$ представляет STDFT или MDCT входного аудио, как определено ранее, STDFT или MDCT обработанного (модифицированного) аудио в Гребенке фильтров синтеза, 110, могут рассчитываться в качестве

$$\hat{X}[k, l] = \sum_b G[b, l] S_b[k] X[k, l - d], \quad (15)$$

где $S_b[k]$ - характеристика фильтра синтеза, ассоциативно связанная с полосой b , а d - задержка, ассоциативно связанная с блоком 109 задержки на фиг.7. Форма фильтров $S_b[k]$ синтеза может выбираться такой же, как фильтры, используемые в гребенке фильтров анализа, $H_b[k]$, или они могут модифицироваться, чтобы обеспечивать безупречную реконструкцию в отсутствии любой модификации коэффициентами усиления (то есть когда $G[b, l]=1$). Конечное обработанное аудио, в таком случае, может формироваться посредством обратного преобразования Фурье или модифицированного косинусного преобразования $X[k, l]$ и синтеза с добавлением перекрытия, как знакомо специалисту в данной области техники.

Целевая удельная громкость

Поведение компоновок, воплощающих аспекты изобретения, таких как примеры по фиг.1-7, главным образом, диктуется образом действий, которым рассчитывается целевая удельная громкость $\hat{N}[b, l]$. Хотя изобретение не ограничено никакими конкретными функцией или обратной функцией для расчета целевой удельной громкости, несколько таких функций и подходящих применений для них будут описаны далее.

Независящая от времени и независящая от частоты функция, пригодная для регулировки громкости

Стандартный регулятор громкости настраивает громкость звукового сигнала применением широкополосного коэффициента усиления к аудио. Обычно коэффициент усиления привязан к ручке или ползунку, которые настраиваются пользователем до тех пор, пока громкость аудио не находится на требуемом уровне. Аспект настоящего изобретения предусматривает более психоакустически совместимый способ реализации такого регулятора. Согласно этому аспекту изобретения предпочтительнее, чем обладание широкополосным коэффициентом усиления, привязанным к регулятору уровня громкости, который имеет результатом изменение коэффициента усиления на одинаковую величину по всем полосам частот, которое может вызывать изменение в воспринимаемом спектре, коэффициент масштабирования удельной громкости взамен ассоциативно связан с настройкой регулятора уровня громкости, с тем чтобы коэффициент усиления на каждой из многочисленных полос частот изменялся на величину, которая учитывает модель человеческого слуха, так что в соответствии с идеалом нет никакого изменения воспринимаемого спектра. В контексте этого аспекта изобретения и примерного его применения «постоянный» или «независящий от времени» подразумеваются предусматривающими изменения в настройке коэффициента масштабирования регулятора уровня громкости, время от времени, например, пользователем. Такая «независимость от времени» иногда указывается ссылкой как «почти независящий от времени», «квазистационарный», «кусочно независящий от времени», «кусочно стационарный», «ступенчато независящий от времени» и «ступенчато стационарный». При условии такого масштабного коэффициента, α , целевая удельная громкость может рассчитываться в качестве измеренной удельной громкости, умноженной на α .

$$\hat{N}[b, t] = \alpha N[b, t]. \quad (16)$$

Так как полная громкость $L[t]$ является суммой удельной громкости $M[b, t]$ по всем полосам b , вышеприведенная модификация также масштабирует полную громкость коэффициентом α , но она делает это способом, который сохраняет одинаковый воспринимаемый спектр в конкретный момент времени для изменений настройки регулятора уровня громкости. Другими словами, в любой конкретный момент времени изменение настройки регулятора уровня громкости имеет результатом изменение воспринимаемой громкости, но никакого изменения в воспринимаемом спектре модифицированного аудио в сравнении с воспринимаемым спектром немодифицированного аудио. Фиг.13а изображает результирующие многополосные коэффициенты усиления $G[b, t]$ по всем полосам « b » в конкретный момент времени « t », когда $\alpha=0,25$ для звукового сигнала, состоящего из женской речи. Для сравнения, также начерчен широкополосный коэффициент усиления, требуемый для масштабирования исходной полной громкости посредством 0,25 (горизонтальная линия), как в стандартном регуляторе уровня громкости. Многополосный коэффициент $G[b, t]$ увеличивается в полосах низких и высоких частот по сравнению с полосами средних частот. Это совместимо с кривыми равной громкости, указывающими, что человеческое ухо менее чувствительно на низких и высоких частотах.

Фиг.13b изображает удельную громкость для исходного звукового сигнала, модифицированного широкополосным коэффициентом усиления сигнала, который модифицировался в соответствии с регулятором уровня громкости предшествующего уровня техники, и модифицированного многополосным коэффициентом усиления

сигнала, который модифицировался в соответствии с этим аспектом изобретения. Удельная громкость модифицированного многополосным коэффициентом усиления сигнала является таковой у исходного сигнала, масштабированной на 0,25. Удельная громкость модифицированного широкополосным коэффициентом усиления сигнала изменила его спектральную форму по отношению к таковой у исходного немодифицированного сигнала. В этом случае удельная громкость, в относительном смысле, теряет громкость как на низких, так и на высоких частотах. Это воспринимается как потускнение аудио по мере того, как понижается его уровень громкости, проблема, которая не возникает при многополосном модифицированном сигнале, чья громкость регулируется коэффициентами усиления, полученными в области громкости восприятия.

Наряду с искажением воспринимаемого спектрального баланса, ассоциативно связанным с традиционным регулятором уровня громкости, существует вторая проблема. Свойство восприятия громкости, которое отражается в модели громкости, отраженной в уравнениях 11a-11d, состоит в том, что громкость сигнала на любой частоте уменьшается быстрее по мере того, как уровень сигнала подходит к порогу слышимости. Как результат электрическое затухание, требуемое для передачи такого же затухания громкости более тихому сигналу, является меньшим, чем таковое, требуемое для более громкого сигнала. Традиционный регулятор уровня громкости придает постоянное затухание независимо от уровня сигнала, а потому тихие сигналы становятся «слишком тихими» по отношению к более громким сигналам по мере того, как уровень громкости убавляется. Во многих случаях это имеет следствием потерю деталей в аудио. Рассмотрим запись кастаньет в реверберирующем помещении. В такой записи основной «удар» кастаньет слишком громок по сравнению с реверберирующими эхо-сигналами, но это реверберирующие эхо-сигналы, которые передают размер помещения. По мере того как уровень громкости убавляется традиционным регулятором уровня громкости, реверберирующие эхо-сигналы становятся тише относительно основного удара и в итоге исчезают ниже порога слышимости, оставляя «сухое» звучание кастаньет. Основанный на громкости регулятор уровня громкости предотвращает исчезновение более тихих частей записей посредством подъема более тихих реверберирующих частей записи относительно более громкого основного удара, так что относительная громкость между этими частями остается постоянной. Для того чтобы достичь этого эффекта, многополосные коэффициенты усиления $G[b, t]$ должны меняться во времени со скоростью, которая соизмерима с человеческим временным разрешением восприятия громкости. Так как многополосные коэффициенты усиления $G[b, t]$ вычисляются в качестве функции сглаженной накачки $\bar{E}[b, t]$, выбор постоянных λ_b времени в уравнении 8 предписывает, насколько быстро коэффициенты усиления могут меняться во времени в каждой полосе b . Как упомянуто ранее, эти постоянные времени могут выбираться пропорциональными времени интегрирования человеческого восприятия громкости в пределах полосы b и, таким образом, давать надлежащее изменение $G[b, t]$ со временем. Должно быть отмечено, что, если временные ограничения выбираются ненадлежащим образом (слишком быстрыми либо слишком медленными), то неприятные по восприятию артефакты могут привноситься в обработанное аудио.

Независящая от времени и зависящая от частоты функция, пригодная для постоянной амплитудной коррекции

В некоторых применениях кто-то может пожелать применять постоянную амплитудную коррекцию восприятия к аудио, в каковом случае целевая удельная

громкость может вычисляться применением независящего от времени, но зависящего от частоты масштабного коэффициента $\Theta[b]$, как в зависимости

$$\hat{N}[b, t] = \Theta[b]N[b, t],$$

в которой $\hat{N}[b, t]$ - целевая удельная громкость, $N[b, t]$ - удельная громкость звукового сигнала, b - показатель частоты, а t - показатель времени. В этом случае масштабирование может изменяться от полосы к полосе. Такое применение, например, может быть полезным для подчеркивания части спектра, доминируемой речевыми частотами для того, чтобы повышать разборчивость.

Независящая от частоты и зависящая от времени функция для автоматической регулировки усиления и регулировки динамического диапазона

Технологии автоматической регулировки усиления и регулировки динамического диапазона (APU и DRC) широко известны в области звуковой обработки. В абстрактном смысле обе технологии измеряют уровень звукового сигнала некоторым образом, а затем модифицируют сигнал коэффициентом усиления на величину, которая является функцией измеренного уровня. Для случая APU сигнал модифицируется коэффициентом усиления, так что его измеренный уровень ближе к выбранному пользователем контрольному уровню. С помощью DRC сигнал модифицируется коэффициентом усиления, так что диапазон измеренного уровня сигнала преобразуется в некоторый требуемый диапазон. Например, кто-то может пожелать сделать тихие части аудио более громкими, а громкие части более тихими. Такая система описана Робинсоном и Гундри (Charles Robinson and Kenneth Gundry, «Dynamic Range Control via Metadata», 107th Convention of the AES, Preprint 5028, September 24-27, 1999, New York (Чарльз Робинсон и Кеннет Гундри, «Регулировка динамического диапазона посредством метаданных», 107^{ая} конвенция AES, препринт 5028, 24-27 сентября 1999 года, Нью-Йорк)). Традиционные реализации APU и DRC обычно используют простое измерение уровня звукового сигнала, такое как сглаженная пиковая или среднеквадратическая (rms) амплитуда, для получения модификации коэффициента усиления. Такие простые измерения до некоторой степени могут коррелировать с воспринимаемой громкостью аудио, но аспекты настоящего изобретения предусматривают более значимые по восприятию APU и DRC посредством модификаций коэффициентом усиления с измерением громкости на основании психоакустической модели. К тому же многие традиционные системы APU и DRC применяют модификацию коэффициента усиления с помощью широкополосного коэффициента усиления, тем самым навлекая на себя вышеупомянутые тембровые (спектральные) искажения в обработанном аудио. Аспекты настоящего изобретения, с другой стороны, используют многополосный коэффициент усиления для придания формы удельной громкости некоторым образом, который снижает или минимизирует такие искажения.

Оба применения, APU и DRC, применяющие аспекты настоящего изобретения, характеризуются функцией, которая преобразует или отображает входную широкополосную громкость $L_i[t]$ в требуемую выходную широкополосную громкость $L_o[t]$, где громкость измеряется в единицах громкости восприятия, таких как соны. Входная широкополосная громкость $L_i[t]$ является функцией удельной громкости $N[b, t]$ входного звукового сигнала. Хотя она может быть такой же, как полная громкость входного звукового сигнала, она может быть сглаженным во времени вариантом полной громкости звукового сигнала.

Фиг.14а и 14б изображает примеры функций отображения, типичных для APU и DRC соответственно. При условии такого отображения, в котором $L_o[t]$ является

функцией от $L_i[t]$, целевая удельная громкость может рассчитываться в качестве

$$\hat{N}[b, t] = \frac{L_o[t]}{L_i[t]} N[b, t]. \quad (17)$$

Исходная удельная громкость $M[b, t]$ звукового сигнала просто масштабируется отношением требуемой выходной широкополосной громкости ко входной широкополосной громкости, чтобы давать выходную удельную громкость $\hat{N}[b, t]$. Для системы АРУ входная широкополосная громкость $L_i[t]$ обычно должна быть показателем долговременной полной громкости аудио. Это может достигаться сглаживанием полной громкости $L[t]$ по времени для формирования $L_i[t]$.

По сравнению с АРУ система DRC реагирует на более кратковременные изменения в громкости сигнала, а потому $L_i[t]$ может просто делаться равной $L[t]$. Как результат масштабирование удельной громкости, заданное посредством $L_o[t]/L_i[t]$, может быстро флуктуировать, приводя к нежелательным артефактам в воспринимаемом аудио. Одним из типичных артефактов является слышимая модуляция части частотного спектра некоторой другой относительно несвязанной частью спектра. Например, подборка классической музыки могла бы содержать высокие частоты, доминируемые продолжительной струнной нотой, наряду с тем, что низкие частоты содержат громко гроыхающие литавры. Всякий раз, когда ударяют литавры, общая громкость $L_i[t]$ повышается, и система DRC применяет затухание к взятой в целом удельной громкости. Струны, в таком случае, слышатся «опустошаемыми» и наполняемыми в зависимости от литавров. Такая перекрестная пульсация в спектре также является проблемой при традиционной системе с широкополосной DRC, и типичное решение влечет за собой применение DRC независимо по отношению к разным полосам частот. Система, раскрытая здесь, по своей природе является многополосной вследствие гребенки фильтров и расчета удельной громкости, которые применяют модель громкости восприятия, а потому модификация системы DRC для работы многополосным образом в соответствии с аспектами настоящего изобретения является относительно простой и описана далее.

Зависящая от времени и зависящая от частоты функция, пригодная для регулировки динамического диапазона

Система DRC может быть расширена для работы многополосным или зависящим от частоты образом предоставлением входной и выходной громкости возможности независимо меняться с полосой b . Эти значения многополосной громкости указываются ссылкой как $L_i[b, t]$ и $L_o[b, t]$, и в таком случае целевая удельная громкость может быть задана согласно

$$\hat{N}[b, t] = \frac{L_o[b, t]}{L_i[b, t]} N[b, t], \quad (18)$$

где $L_o[b, t]$ была рассчитана по или отображена из $L_i[b, t]$, как проиллюстрировано на фиг. 14b, но независимо для каждой полосы b . Входная многополосная громкость $L_i[b, t]$ является функцией удельной громкости $M[b, t]$ входного звукового сигнала. Хотя она может быть такой же, как удельная громкость входного звукового сигнала, она может быть сглаженным во времени и/или сглаженным по частоте вариантом удельной громкости звукового сигнала.

Наиболее простой способ расчета $L_i[b, t]$ состоит в том, чтобы устанавливать ее равной удельной громкости $M[b, t]$. В этом случае DRC скорее выполняется независимо в каждой полосе в гребенке слуховых фильтров модели громкости восприятия, чем в соответствии с одинаковым отношением входной к выходной громкости для всех полос, как только что описано выше под заголовком «Независящая от частоты и

зависящая от времени функция, пригодная для автоматической регулировки усиления и регулировки динамического диапазона». В практическом варианте осуществления, применяющем 40 полос, разнесение этих полос по оси частот является относительно частым, для того чтобы давать точный показатель громкости. Однако применение

5 масштабного коэффициента DRC независимо к каждой полосе может заставлять обработанное аудио звучать «раздробленно». Чтобы избежать этой проблемы, можно предпочесть рассчитывать $L_i[b, t]$ сглаживанием удельной громкости $M[b, t]$ по

10 полосам, так что применяемая величина DRC от одной полосы к следующей не меняется настолько радикально. Это может достигаться определением фильтра $Q(b)$ сглаживания полос, а затем сглаживанием удельной громкости по всем полосам с согласно стандартной сверточной сумме:

$$L_i[b, t] = \sum_c Q(b-c) M[c, t]. \quad (19)$$

15 при этом $M[c, t]$ - удельная громкость звукового сигнала, а $Q(b-c)$ - сдвигаемая по полосам характеристика сглаживающего фильтра. Фиг.15 изображает один из примеров такого сглаживающего полосы фильтра.

Если функция DRC, которая рассчитывает $L_i[b, t]$ в качестве функции от $L_o[b, t]$, является фиксированной для каждой полосы b , то тип изменения, вносимого в каждую

20 полосу удельной громкости $M[b, t]$, будет меняться в зависимости от спектра аудио, являющегося обрабатываемым, даже если общая громкость сигнала остается той же самой. Например, звуковой сигнал с громким басом и тихим сопрано может содержать бас срезанным, а сопрано повышенным. Сигнал с тихим басом и громким

25 сопрано может заставлять происходить противоположное. Результирующим эффектом является изменение тембра или воспринимаемого спектра аудио, а это может быть желательным в определенных применениях.

Однако кто-то может пожелать выполнять многополосную DRC, не модифицируя нормальный воспринимаемый спектр аудио. Кто-то мог пожелать, чтобы средняя

30 модификация в каждой полосе была приблизительно одинаковой наряду с прежним предоставлением возможности кратковременных изменений модификаций для независимой работы между или среди полос. Желательный эффект может достигаться вынуждением среднего поведения DRC в каждой полосе быть таким же, как таковое у

35 некоторого среднего эталонного поведения. Можно выбирать это эталонное поведение в качестве требуемого DRC для широкополосной входной громкости $L_i[t]$. Пусть функция $L_o[t] = DRC\{L_i[t]\}$ представляет требуемое отображение DRC для широкополосной громкости. В таком случае пусть $\bar{L}[b, t]$ представляет усредненный по времени вариант многополосной входной громкости $L_i[b, t]$. Многополосная

40 выходная громкость в таком случае может рассчитываться в качестве

$$L_o[b, t] = \frac{\bar{L}_i[b, t]}{\bar{L}_i[t]} DRC\left\{ \frac{\bar{L}_i[t]}{\bar{L}_i[b, t]} L_i[b, t] \right\}. \quad (20)$$

Отметим, что многополосная входная громкость сначала масштабируется, чтобы

45 быть в таком же среднем диапазоне, как широкополосная входная громкость. Затем применяется функция DRC, предназначенная для широкополосной громкости. Наконец, результат подвергается уменьшению масштаба обратно, в средний диапазон многополосной громкости. С помощью этой формулировки многополосной DRC

50 сохраняется пониженная спектральная пульсация наряду с одновременным сохранением среднего воспринимаемого спектра аудио.

Зависящая от частоты и зависящая от времени функция, пригодная для динамической коррекции

Еще одним применением аспектов настоящего изобретения является преднамеренная трансформация зависящего от времени воспринимаемого спектра аудио в целевой независимый от времени спектр, по-прежнему наряду с сохранением исходного динамического диапазона аудио. На эту обработку могут указывать ссылкой как динамическую коррекцию (DEQ). При традиционной статической коррекции простая постоянная фильтрация применяется к аудио, для того чтобы изменять его спектр. Например, можно применять постоянный подъем баса и сопрано. Такая обработка не принимает во внимание текущий спектр аудио, а потому может быть неподходящей для некоторых сигналов, то есть сигналов, которые уже содержат относительно большую величину баса или сопрано. При DEQ спектр сигнала измеряется, и сигнал затем динамически модифицируется, для того чтобы трансформировать измеренный спектр в по существу статическую требуемую форму. Что касается аспектов настоящего изобретения, такая требуемая форма задается по полосам в гребенке фильтров и указывается ссылкой как $EQ[b]$. В практическом варианте осуществления измеренный спектр должен представлять среднюю спектральную форму аудио, которая может формироваться сглаживанием удельной громкости $M[b, t]$ по времени. Сглаженную удельную громкость могут указывать ссылкой как $\bar{M}[b, t]$. Как при многополосной DRC, кто-то может не пожелать, чтобы модификация DEQ радикально менялась от одной полосы к следующей, а потому функция сглаживания полос может применяться для формирования сглаженного по полосам спектра $\bar{L}[b, t]$:

$$\bar{L}[b, t] = \sum_c Q(b - c) \bar{M}[c, t]. \quad (21)$$

Для того чтобы сохранить исходный динамический диапазон аудио, требуемый спектр $EQ[b]$ должен быть нормализован, чтобы иметь такую же общую громкость, как измеренная спектральная форма, заданная посредством $\bar{L}[b, t]$. Эту нормализованную спектральную форму могут указывать ссылкой как $\bar{L}_{EQ}[b, t]$:

$$\bar{L}_{EQ}[b, t] = \left(\frac{\sum_c \bar{L}[c, t]}{\sum_c EQ[c]} \right) EQ[b]. \quad (22)$$

В заключение целевая удельная громкость рассчитывается в качестве

$$\hat{M}[b, t] = \left(\frac{\bar{L}[b, t]}{\bar{L}_{EQ}[b, t]} \right)^{1-\beta} \frac{\bar{L}_{EQ}[b, t]}{\bar{L}[b, t]} M[b, t], \quad (23)$$

где β - заданный пользователем параметр, находящийся в диапазоне от нуля до единицы, указывающий степень DEQ, которая должна применяться. Глядя на уравнение 23, кто-то заметит, что когда $\beta=0$, исходная удельная громкость является немодифицированной, а когда $\beta=1$, удельная громкость масштабируется отношением требуемой спектральной формы к измеренной спектральной форме.

Один из удобных способов формирования требуемой спектральной формы $EQ[b]$ предназначен, чтобы пользователь устанавливал ее равной $\bar{L}[b, t]$, которая измерена для некоторой части аудио, чей спектральный баланс находит приятным пользователь. В практическом варианте осуществления, например, как показанный на фиг.16, пользователь может снабжаться кнопкой или другим подходящим исполнительным механизмом 507, который, когда приведен в действие, вызывает фиксацию текущего показателя спектральной формы $\bar{L}[b, t]$ аудио, а затем сохраняет этот показатель в качестве предустановки (в Фиксации и хранении предустановки целевой удельной громкости, 506), которая позже может загружаться в $EQ[b]$, когда задействована DEQ (как посредством выбора 508 предустановки). Фиг.16 -

упрощенный вариант фиг.7, в котором только одна линия показана для представления многочисленных полос из Гребенки фильтров анализа, 100, в Гребенку фильтров синтеза, 110. Пример фиг.17 также предусматривает Модификацию удельной громкости (SL) динамической EQ, 505, которая обеспечивает модификацию в отношении удельной громкости, измеренной функцией или устройством 104, в соответствии с динамической коррекцией, как поясненная выше.

Комбинированная обработка

Кто-то может пожелать объединить всю описанную ранее обработку, включая регулировку уровня громкости (VC), APY, DRC и DEQ в единую систему. Так как каждая из этих последовательностей операций может быть представлена в качестве масштабирования удельной громкости, все из них легко объединяются, как изложено ниже:

$$\hat{M}[b, t] = (\Xi_{VC}[b, t] \Xi_{AGC}[b, t] \Xi_{DRC}[b, t] \Xi_{DEQ}[b, t]) M[b, t], \quad (24)$$

где $\Xi_*[b, t]$ представляет масштабные коэффициенты, ассоциативно связанные с последовательностью операций «*». Единый набор коэффициентов $G[b, t]$ усиления затем может рассчитываться для целевой удельной громкости, который представляет комбинированную обработку.

В некоторых случаях масштабные коэффициенты одной или комбинации последовательностей операций модификации громкости могут флуктуировать слишком быстро со временем и создавать артефакты в результирующем обработанном аудио. Поэтому может быть желательным сглаживать некоторые подмножества этих коэффициентов масштабирования. Вообще, масштабные коэффициенты из VC и DEQ равномерно меняются со временем, но может потребоваться сглаживание комбинации масштабных коэффициентов APY и DRC. Пусть комбинация этих масштабных коэффициентов представлена посредством

$$\Xi_C[b, t] = \Xi_{AGC}[b, t] \Xi_{DRC}[b, t] \quad (25)$$

Основная идея за сглаживанием состоит в том, что комбинированные масштабные коэффициенты должны быстро реагировать, когда удельная громкость повышается, и что масштабные коэффициенты должны сильнее сглаживаться, когда удельная громкость уменьшается. Эта идея соответствует широко известной практике использования быстрого наступления и медленного освобождения в конструкции звуковых компрессоров. Надлежащие постоянные времени для сглаживания масштабных коэффициентов могут рассчитываться сглаживанием по времени сглаженного по полосам варианта удельной громкости. Прежде всего вычисляется сглаженный по полосам вариант удельной громкости:

$$L[b, t] = \sum_c Q(b-c) M[c, t]. \quad (26)$$

при этом $M[c, t]$ - удельная громкость звукового сигнала, а $Q(b-c)$ - сдвигаемая по полосам характеристика сглаживающего фильтра, как в уравнении 19, приведенном выше.

Сглаженный по времени вариант этой сглаженной по полосам удельной громкости затем рассчитывается в качестве

$$\bar{L}[b, t] = \lambda[b, t] L[b, t] + (1 - \lambda[b, t]) \bar{L}[b, t-1] \quad (27)$$

где зависимый от полосы коэффициент сглаживания $\lambda[b, t]$ задан согласно

$$\lambda[b, t] = \begin{cases} \lambda_{fast}, & L[b, t] > \bar{L}[b, t] \\ \lambda_{slow}, & L[b, t] \leq \bar{L}[b, t] \end{cases} \quad (28)$$

Сглаженные комбинированные масштабные коэффициенты затем рассчитываются как

$$\Xi_c[b, t] = \lambda_M[b, t] \Xi_c[b, t] + (1 - \lambda_M[b, t]) \Xi_c[b, t-1], \quad (29)$$

где $\lambda_M[b, t]$ - сглаженный по полосам вариант $\lambda[b, t]$:

$$\lambda_M[b, t] = \left(\frac{1}{\sum_c Q(c)} \right) \sum_c Q(b-c) \lambda[b, t]. \quad (30)$$

Сглаживание по полосам коэффициентов сглаживания предохраняет сглаженные по времени масштабные коэффициенты от радикального изменения по полосам.

Описанное сглаживание по времени и полосам масштабных коэффициентов имеет результатом обработанное аудио, содержащее меньшее количество нежелательных относящихся к восприятию артефактов.

Компенсация шума

Во многих средах воспроизведения аудио существует фоновый шум, который мешает аудио, которое желает прослушивать слушатель. Например, слушатель в движущемся автомобиле может быть проигрывающим музыку через встроенную стереофоническую систему, и шум от двигателя и дороги может значительно изменять восприятие музыки. В частности, для частей спектра, в которых энергия шума значительна относительно энергии музыки, воспринимаемая громкость музыки снижается. Если уровень шума достаточно высок, музыка полностью скрывается. Что касается аспекта настоящего изобретения, кто-то захотел бы выбрать коэффициенты $G[b, t]$ усиления, с тем чтобы удельная громкость обработанного аудио в присутствии мешающего шума была равной целевой удельной громкости $\hat{N}[b, t]$. Для достижения этого результата можно использовать концепцию громкости частичных тонов, как определено Муром и Глазбергом ранее. Допустим, что кто-то способен получать измерение шума самого по себе и измерение аудио самого по себе. Пусть $E_N[b, t]$ представляет накачку от шума, и пусть $E_A[b, t]$ представляет накачку от аудио. Объединенная удельная громкость аудио и шума, в таком случае, задается посредством

$$N_{tot}[b, t] = \Psi\{E_A[b, t] + E_N[b, t]\}, \quad (31)$$

где снова $\Psi\{\cdot\}$ представляет нелинейное преобразование из накачки в удельную громкость. Можно допустить, что слух слушателя разделяет объединенную удельную громкость между удельной громкостью частичных тонов аудио и удельной громкостью частичных тонов шума некоторым способом, который сохраняет объединенную удельную громкость:

$$N_{tot}[b, t] = N_A[b, t] + N_N[b, t]. \quad (32)$$

Удельная громкость частичных тонов аудио, $N_A[b, t]$, является значением, которое желательно контролировать, а потому необходимо вычислять это значение. Удельная громкость частичных тонов шума может быть приближенно выражена в качестве

$$N_N[b, t] = \left(\frac{E_{TN}[b, t]}{E_A[b, t]} \right)^k \left(\Psi\{E_N[b, t] + E_{TN}[b, t]\} - \Psi\{E_{TQ}[b]\} \right) \quad (33)$$

где $E_{TN}[b, t]$ - маскируемое пороговое значение в присутствии шума, $E_{TQ}[b]$ - порог слышимости в тишине на полосе b , и k - показатель степени между нулем и единицей.

Объединяя уравнения 31-33, приходим к выражению для удельной громкости частичных тонов аудио:

$$N_A[b, t] = \Psi\{E_A[b, t] + E_N[b, t]\} - \left(\frac{E_{TN}[b, t]}{E_A[b, t]} \right)^k \left(\Psi\{E_N[b, t] + E_{TN}[b, t]\} - \Psi\{E_{TQ}[b]\} \right) \quad (34)$$

Отмечают, что, когда накачка аудио равна маскируемому пороговому значению шума ($E_A[b, t] = E_{TN}[b, t]$), удельная громкость частичных тонов аудио равна

громкости сигнала на пороговом значении в тишине, каковое является требуемым результатом. Когда накачка аудио гораздо больше, чем у шума, второй член в уравнении 34 исчезает, и удельная громкость аудио приблизительно равна такой, какая могла бы быть, если бы шум отсутствовал. Другими словами, в то время как аудио становится гораздо громче чем шум, шум скрывается за аудио. Показатель k выбирается опытным путем, чтобы давать максимальное соответствие данным о громкости тона в шуме в качестве функции отношения сигнал-шум. Мур и Глазберг обнаружили, что значение $k=0,3$ является подходящим. Маскируемое пороговое значение шума может быть приближенно выражено в качестве функции восприятия собственно шума:

$$E_{TN}[b, t] = K[b]E_N[b, t] + E_{TQ}[b] \quad (35)$$

где $K[b]$ - константа, которая увеличивается в полосах нижних частот. Таким образом, удельная громкость частичных тонов аудио, заданная уравнением 34, может быть представлена абстрактно, в качестве функции накачки аудио и накачки шума:

$$N_A[b, t] = \Phi\{E_A[b, t], E_N[b, t]\}. \quad (36)$$

Модифицированный решатель коэффициентов усиления в таком случае может использоваться для расчета коэффициентов усиления $G[b, t]$ из условия, чтобы удельная громкость частичных тонов обработанного аудио в присутствии шума была равна целевой удельной громкости:

$$\hat{N}[b, t] = \Phi\{G^2[b, t]E_A[b, t], E_N[b, t]\} \quad (37)$$

Фиг.17 изображает систему по фиг.7 с исходным Решателем коэффициентов усиления, 106 замещенным описанным Решателем коэффициентов усиления с компенсацией шума, 206, (отметим, что многочисленные вертикальные линии между блоками, представляющими многочисленные полосы гребенки фильтров, были замещены одиночной линией для упрощения схемы). В дополнение фигура изображает измерение накачки шума (Гребенкой фильтров анализа, 200, Пропускающим фильтром 201, Накачкой, 202, и Сглаживанием, 203, некоторым образом, соответствующим работе блоков 100, 101, 102 и 103), которое подается в новый решатель 206 коэффициентов усиления наряду с накачкой аудио (из Сглаживания, 103) и целевой удельной громкостью (из Модификации SL, 105).

В своем наиболее основном режиме работы Модификация SL, 105, по фиг.17 может просто устанавливать целевую удельную громкость $\hat{N}[b, t]$ равной исходной удельной громкости аудио, $M[b, t]$. Другими словами, Модификация SL предусматривает независимый от частоты масштабный коэффициент α , масштабирующий удельную громкость звукового сигнала, при этом $\alpha=1$. С помощью компоновки, такой как на фиг.17, коэффициенты усиления рассчитываются так, что воспринимаемый спектр громкости обработанного аудио в присутствии шума равен спектру громкости аудио в отсутствие шума. Дополнительно, любая одна или комбинация таковых из описанных ранее технологий для вычисления целевой удельной громкости в качестве функции исходной, в том числе VC, APY, DRC и DEQ, могут использоваться в соединении с системой модификации громкости с компенсацией шума.

В практическом варианте осуществления измерение шума может быть получено с микрофона, размещенного в или возле среды, в которую будет воспроизводиться аудио. В качестве альтернативы может использоваться предопределенный набор шаблонных накачек шума, которые аппроксимируют ожидаемый спектр шума при различных условиях. Например, шум в салоне автомобиля может предварительно анализироваться при различных скоростях езды, а затем сохраняться в качестве справочной таблицы шумовой накачки в зависимости от скорости. Шумовая накачка,

подаваемая в Решатель коэффициентов усиления, 206, по фиг.17, в таком случае, приближенно выражается по этой справочной таблице, по мере того как изменяется скорость автомобиля.

Приближения к удельной громкости

Несмотря на то что раскрытое изобретение работает наилучшим образом, когда используется точное измерение удельной громкости, некоторые применения могут требовать использования более грубой аппроксимации, для того чтобы снижать вычислительную сложность. С подходящим приближением по-прежнему может достигаться приемлемая оценка и модификация воспринимаемой громкости. Такое приближение должно пытаться сохранить, по меньшей мере частично, несколько ключевых аспектов восприятия громкости. Во-первых, приближение должно, по меньшей мере грубо, фиксировать изменение чувствительности в восприятии громкости в зависимости от частоты. В общих чертах, приближение должно отображать меньшую чувствительность на нижних и верхних частотах по сравнению со средними частотами. Во-вторых, приближение должно демонстрировать нелинейный рост громкости в зависимости от уровня сигнала. Более точно, рост удельной громкости должен быть наиболее быстрым для низкоуровневых сигналов возле порога слышимости, а затем уменьшаться до постоянной скорости роста по мере того, как повышается уровень сигнала. Наконец, приближение должно демонстрировать свойство суммирования громкости, означающее, что для постоянного уровня сигнала полная громкость (интегрирование удельной громкости по частоте) увеличивается по мере того, как увеличивается ширина полосы пропускания сигнала.

Один из способов для уменьшения сложности вычисления удельной громкости, по-прежнему наряду с сохранением желательных свойств восприятия громкости, состоит в том, чтобы использовать гребенку фильтров с меньшим количеством полос, и в которой полосы могут не быть равномерно разнесенными по ступенчатой шкале критических полос. Например, можно использовать 5-полосную гребенку фильтров с полосами, равномерно распределенными по линейной шкале частот, в противоположность 40 полосам, описанным ранее. Существует много технологий для эффективной реализации таких гребенок фильтров, например, модулированные косинусом гребенки фильтров (P.P Vaidyanthan, *MultiRate Systems and Filter Banks*, 1993 Prentice Hall (П. П. Вайдьянтан, *Многоступенчатые системы и гребенки фильтров*, 1993 год, Прентис-Холл)). В качестве общего примера рассмотрим гребенку фильтров с B полосами, где каждая полоса описывается импульсной характеристикой $h_b[n]$ во временной области. К тому же, допустим, что гребенка фильтров является почти совершенной реконструкцией, имея в виду, что:

$$x[n] \approx \sum_{b=1}^B h_b[n] * x[n] \quad (38)$$

Частотная характеристика каждой полосы b может характеризоваться центральной частотой f_b и шириной Δf_b полосы пропускания в Герцах. Эквивалентная центральная частота и ширина полосы пропускания в единицах ERB, в таком случае, задаются согласно

$$e_b = \text{HzToERB}(f_b) \quad (39a)$$

$$\Delta e_b = \text{HzToERB}\left(f_b + \frac{\Delta f_b}{2}\right) - \text{HzToERB}\left(f_b - \frac{\Delta f_b}{2}\right) \quad (39b)$$

Если количество полос B относительно мало, то ширина e_b полосы пропускания каждой полосы, вероятно, будет большей, чем 1 ERB.

С допущением, что $x_b[n] = h_b[n] * x[n]$ представляет звуковой сигнал, ассоциативно связанный с каждой полосой, сглаженный сигнал накачки $\bar{E}[b, n]$ может вычисляться подобно уравнениям 7 и 8, с вычислением сглаженной среднеквадратической мощности $x_b[n]$, взвешенной частотной характеристикой пропускающего фильтра, дискретизированного на частоте f_b , и инверсией ширины полосы пропускания в единицах ERB:

$$\bar{E}[b, n] = \lambda_b \bar{E}[b, n] + (1 - \lambda_b) \frac{|P(f_b)|^2}{e_b} x_b^2[n] \quad (40)$$

Взвешивание накачки полосы b посредством $1/e_b$ эффективно распределяет энергию в пределах такой полосы, равномерно по всем критическим полосам, отнесенным к группе в пределах нее. В качестве альтернативы кто-то мог бы назначать всю энергию на критическую полосу, чья центральная частота наиболее близка к центральной частоте f_b полосы, но распределение энергии равномерно является лучшим приближением для большинства реальных звуковых сигналов.

С накачкой $\bar{E}[b, n]$, можно приступить к вычислению удельной громкости $M[b, n]$, целевой удельной громкости $\hat{N}[b, n]$ и коэффициентов усиления $G[b, n]$, как обсуждено выше, за исключением того, что здесь общее количество полос может быть гораздо меньшим, тем самым снижая сложность. К тому же, одной модификации необходимо производиться в вычислении общей громкости $L[n]$: удельная громкость должна взвешиваться количеством ERB в пределах полосы b , в то время как она суммируется по полосам:

$$L[n] = \sum_b e_b N[b, n] \quad (41)$$

В заключение модифицированный сигнал $y[n]$ может формироваться суммированием каждого из полосных сигналов, взвешенных коэффициентом усиления из соответственной полосы:

$$y[n] = \sum_b G[b, n] x_b[n] \quad (42)$$

Для простоты описания вышеприведенные вычисления показаны выполняемыми для каждого периода n выборки сигнала $x[n]$. На практике, однако, накачка может быть подвергнута понижающей дискретизации по времени до гораздо меньшей частоты, а затем вся последующая обработка громкости может выполняться на этой пониженной частоте. Когда коэффициенты усиления применяются в конце, они затем могут быть подвергнуты повышающей дискретизации посредством интерполяции до применения к полосовым сигналам.

Вышеприведенное является только одним из примеров грубого приближения к удельной громкости, которое является подходящим для раскрытого изобретения. Возможны другие приближения, и изобретение подразумевается покрывающим использование всех таких приближений.

Реализация

Изобретение может быть реализовано в аппаратных средствах или программном обеспечении либо сочетании обоих (например, программируемых логических матрицах). Если не указан иной способ действий, алгоритмы, включенные в качестве части изобретения, по своей природе не имеют отношения к какому бы то ни было конкретному компьютеру или другому устройству. В частности, различные машины общего применения могут использоваться с программами, написанными в соответствии с доктринами, приведенными в материалах настоящей заявки, или может быть более удобным сконструировать более специализированное устройство (например, интегральные схемы) для выполнения требуемых этапов способа. Таким

образом, изобретение может быть реализовано в одной или более компьютерных программ, выполняющихся в одной или более программируемых компьютерных системах, каждая из которых содержит по меньшей мере один процессор, по меньшей мере одну систему хранения данных (в том числе энергозависимую и энергонезависимую память и/или запоминающие элементы), по меньшей мере одно устройство или порт ввода и по меньшей мере одно устройство или порт вывода. Управляющая программа применяется к входным данным для выполнения функций, описанных в материалах настоящей заявки, и формирует выходную информацию. Выходная информация подводится в одно или более устройств вывода известным образом.

Каждая такая программа может быть реализована на любом желательном компьютерном языке (включая машинные, компоновочные или высокоуровневые процедурные, логические или объектно-ориентированные языки программирования) для общения с компьютерной системой. В любом случае, язык может быть компилируемым или интерпретируемым языком.

Каждая такая компьютерная программа предпочтительно хранится на или загружается в запоминающие носители или устройство (например, твердотельную память или носители, либо магнитные или оптические носители), удобочитаемые программируемым компьютером общего применения или специального назначения, для конфигурирования и управления компьютером, когда запоминающие носители или устройства считываются компьютерной системой, чтобы выполнять процедуры, описанные в материалах настоящей заявки. Обладающая признаками изобретения система также может считаться реализуемой в качестве машиночитаемого запоминающего носителя, сконфигурированного компьютерной программой, где запоминающий носитель, сконфигурированный таким образом, побуждает компьютерную систему работать специальным и predetermined образом для выполнения функций, описанных в материалах настоящей заявки.

Было описано некоторое количество вариантов осуществления изобретения. Тем не менее, будет понятно, что различные модификации могут быть произведены, не выходя из сущности и объема изобретения. Например, некоторые из этапов, описанных в материалах настоящей заявки, могут быть независимыми от очередности и, таким образом, могут выполняться в очередности, отличной от той, которая описана.

Формула изобретения

1. Способ для регулировки конкретной характеристики громкости звукового сигнала, в котором конкретная характеристика громкости является либо приближением удельной громкости, которое является показателем громкости восприятия как функции частоты и времени, либо приближением удельной громкости частичных тонов, которое является показателем громкости восприятия сигнала в присутствии вторичного мешающего сигнала, как функции частоты и времени, состоящий в том, что

получают приближение к целевой удельной громкости, получают зависящие от частоты параметры модификации, используемые для модифицирования звукового сигнала, для того, чтобы уменьшать разность между его конкретной характеристикой громкости и приближением к целевой удельной громкости, и

а) применяют параметры модификации к звуковому сигналу для уменьшения разницы между его конкретной характеристикой громкости и приближением к целевой

удельной громкости, либо

б) передают или хранят параметры модификации и звуковой сигнал для разделенного временным образом и/или разделенного пространственным образом применения параметров модификации к звуковому сигналу, чтобы уменьшать разность между его конкретной характеристикой громкости и приближением к целевой удельной громкости.

2. Способ по п.1, в котором приближение к целевой удельной громкости не является функцией звукового сигнала.

3. Способ по п.2, в котором упомянутое модифицирование или упомянутое получение заключается в том, что принимают приближение к целевой удельной громкости из источника, внешнего по отношению к способу.

4. Способ по п.2, в котором упомянутое модифицирование или упомянутое получение включают в себя обработку, которая явным образом рассчитывает приближение к удельной громкости и/или приближение к удельной громкости частичных тонов.

5. Способ по п.2, в котором упомянутое модифицирование или упомянутое получение включают в себя обработку, которая неявным образом рассчитывает приближение к удельной громкости и/или приближение к удельной громкости частичных тонов.

6. Способ по п.5, в котором обработка применяет справочную таблицу из условия, чтобы обработка, по своей природе, определяла приближение к удельной громкости и/или приближение к удельной громкости частичных тонов.

7. Способ по п.5, в котором приближение к удельной громкости и/или приближение к удельной громкости частичных тонов, по своей природе, определяется отражающим ряд решений математическим выражением, применяемым обработкой.

8. Способ по пп.2-7, в котором приближение к целевой удельной громкости является независимым от времени и частоты.

9. Способ по пп.2-7, в котором приближение к целевой удельной громкости является независимым от времени.

10. Способ по п.1, в котором приближение к целевой удельной громкости является функцией звукового сигнала или показателя звукового сигнала.

11. Способ по п.10, в котором показателем звукового сигнала является приближение к удельной громкости звукового сигнала.

12. Способ по п.10 или 11, в котором функцией звукового сигнала или показателя звукового сигнала является одно или более масштабирований звукового сигнала или показателя звукового сигнала.

13. Способ для регулирования конкретной характеристики громкости звукового сигнала, в котором конкретная характеристика громкости является либо приближением удельной громкости, которое является показателем громкости восприятия как функции частоты и времени, либо приближением удельной громкости частичных тонов, которое является показателем громкости восприятия звукового сигнала в присутствии вторичного мешающего сигнала как функции частоты и времени, состоящий в том, что

принимают из передаваемых данных или воспроизводят с запоминающего носителя звуковой сигнал и

а) зависящие от частоты параметры модификации для модифицирования звукового сигнала, параметры модификации были получены из приближения к целевой удельной громкости, или

б) приближение к целевой удельной громкости или представление приближения к такой целевой удельной громкости, и

модифицируют звуковой сигнал в ответ на а) принятые параметры модификации, либо б) параметры модификации, полученные из приближения к целевой удельной громкости или ее представления, для того чтобы уменьшать разность между конкретной характеристикой громкости звукового сигнала и приближением к целевой удельной громкости.

14. Способ по п.13, в котором приближение к целевой удельной громкости не является функцией звукового сигнала.

15. Способ по п.14, в котором упомянутое модифицирование или упомянутое получение заключается в том, что принимают приближение к целевой удельной громкости из источника, внешнего по отношению к способу.

16. Способ по п.14, в котором упомянутое модифицирование или упомянутое получение включают в себя обработку, которая явным образом рассчитывает приближение к удельной громкости и/или приближение к удельной громкости частичных тонов.

17. Способ по п.14, в котором упомянутое модифицирование или упомянутое получение включают в себя обработку, которая неявным образом рассчитывает приближение к удельной громкости и/или приближение к удельной громкости частичных тонов.

18. Способ по п.17, в котором обработка применяет справочную таблицу из условия, чтобы обработка, по своей природе, определяла приближение к удельной громкости и/или приближение к удельной громкости частичных тонов.

19. Способ по п.17, в котором приближение к удельной громкости и/или приближение к удельной громкости частичных тонов, по своей природе, определяется отражающим ряд решений математическим выражением, применяемым обработкой.

20. Способ по любому одному из пп.14-19, в котором приближение к целевой удельной громкости является независимым от времени и частоты.

21. Способ по любому одному из пп.14-19, в котором приближение к целевой удельной громкости является независимым от времени.

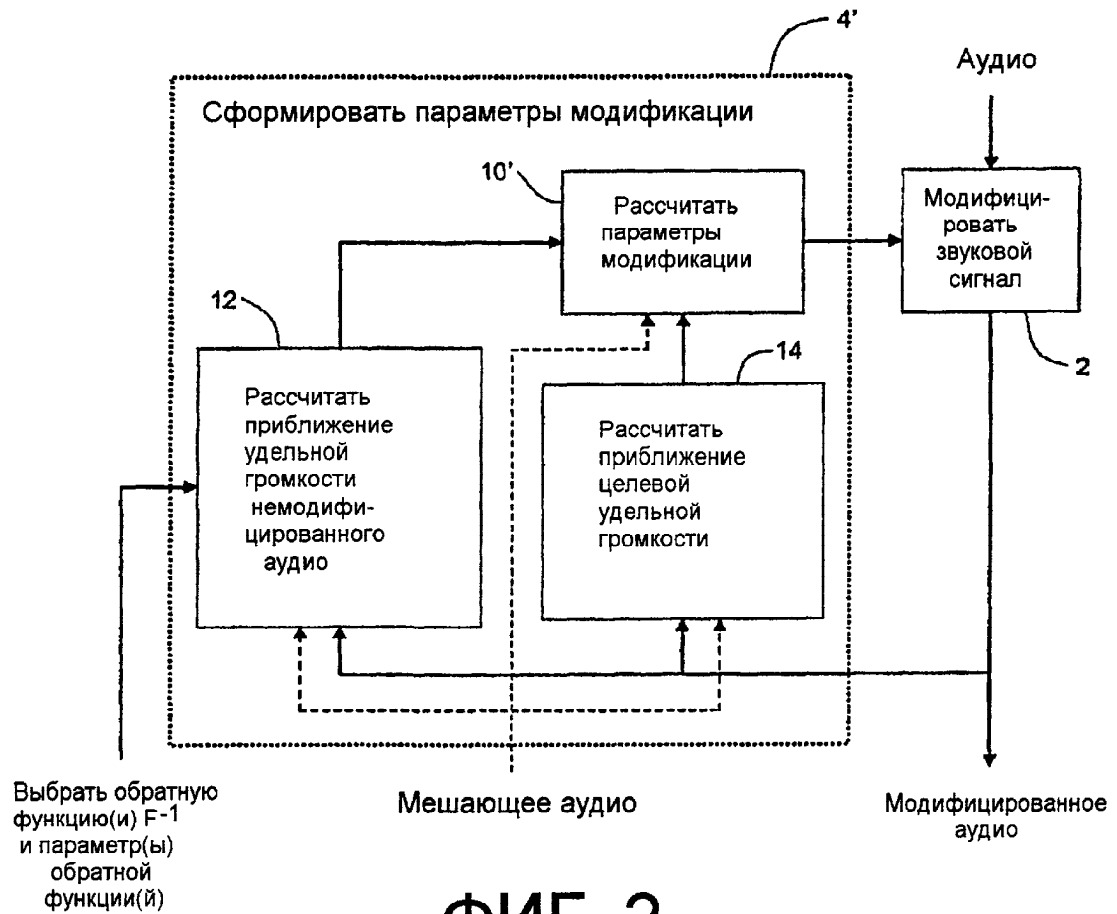
22. Способ по п.13, в котором приближение к целевой удельной громкости является функцией звукового сигнала или показателя звукового сигнала.

23. Способ по п.22, в котором показателем звукового сигнала является приближение к удельной громкости звукового сигнала.

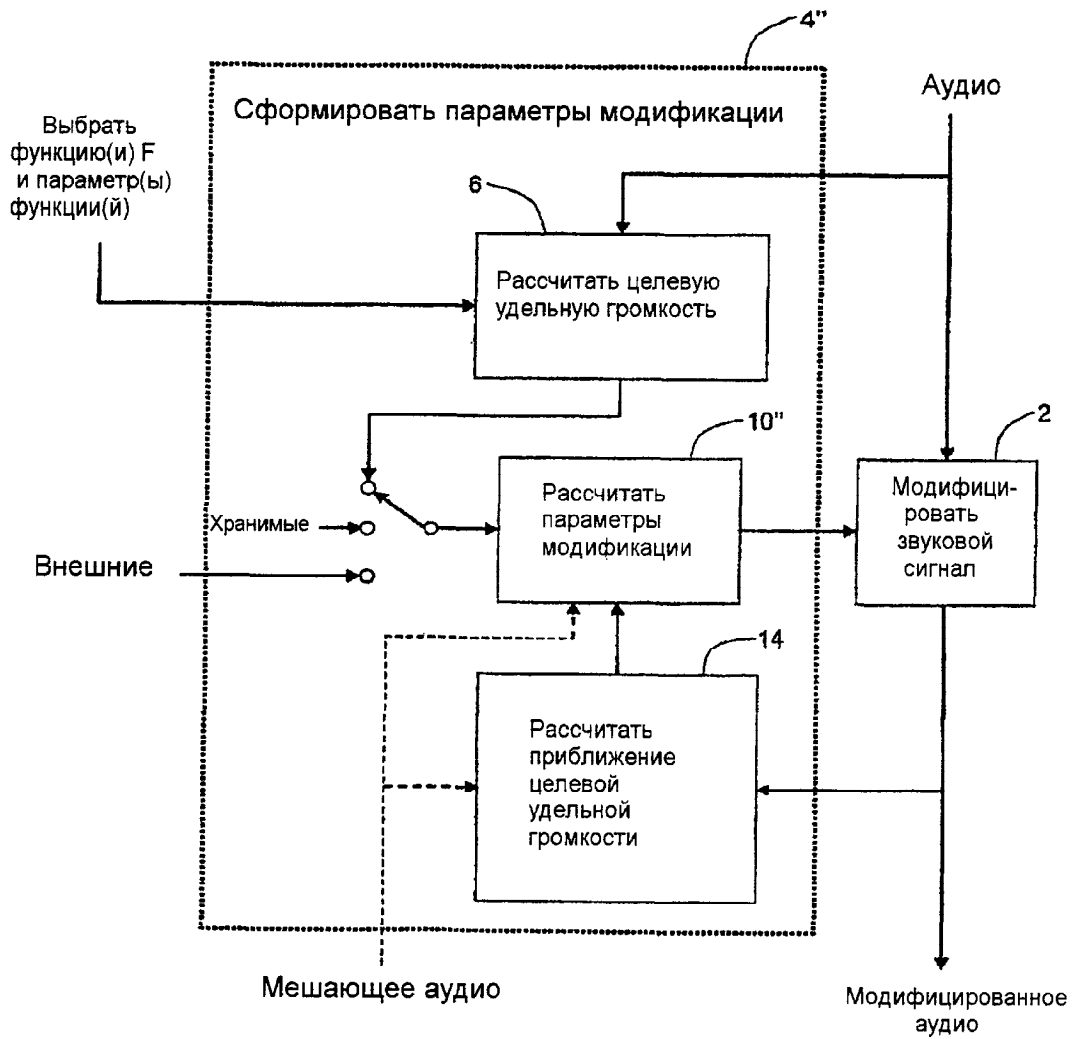
24. Способ по п.22 или 23, в котором функцией звукового сигнала или показателя звукового сигнала является одно или более масштабирований звукового сигнала или показателя звукового сигнала.

25. Устройство для регулирования конкретной характеристики громкости звукового сигнала, выполненное с возможностью осуществления всех этапов способа по любому одному из пп.1-24.

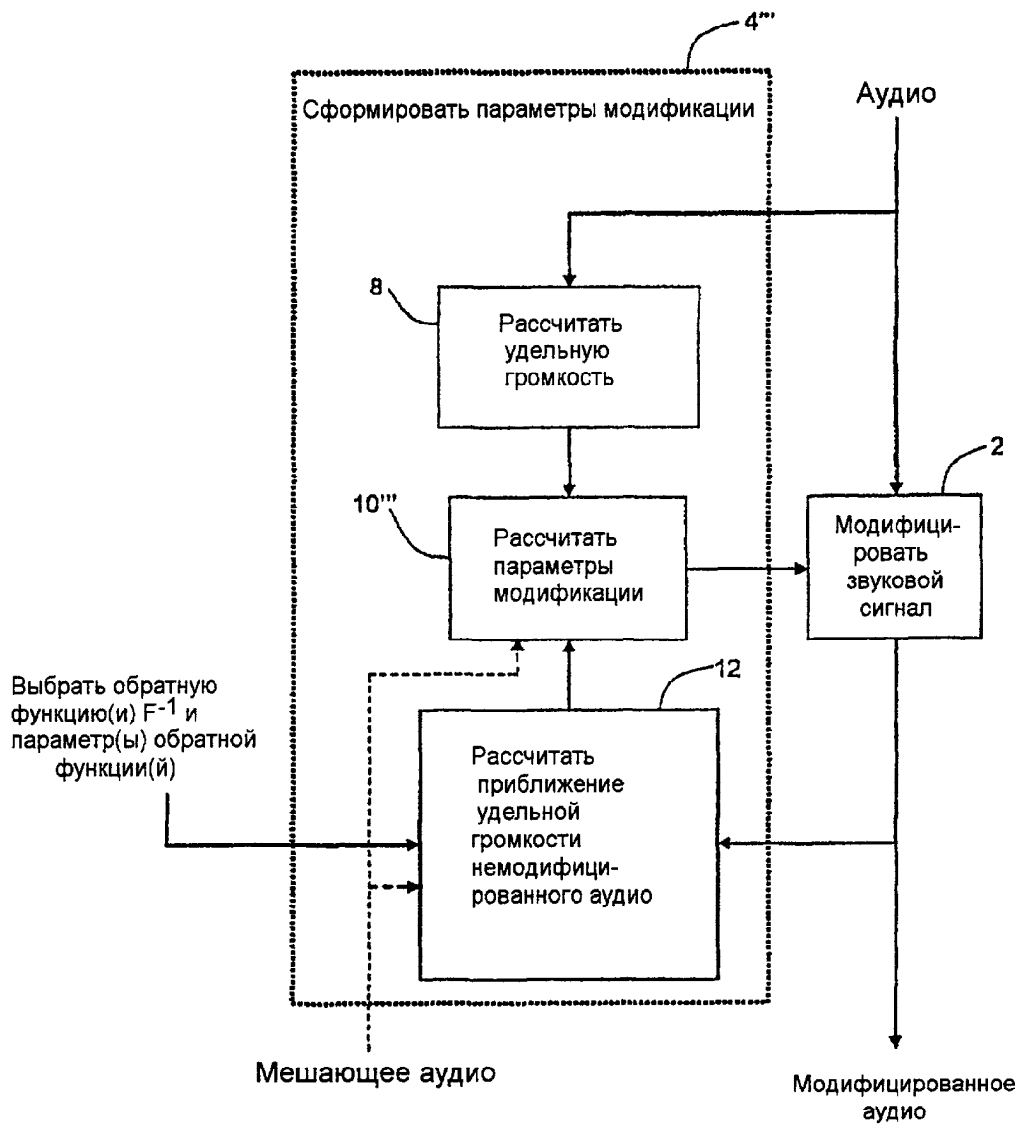
26. Машиночитаемый носитель, содержащий записанную на нем компьютерную программу, при этом компьютерная программа при исполнении ее компьютером приводит к осуществлению упомянутым компьютером всех этапов способа по любому одному из пп.1-24.



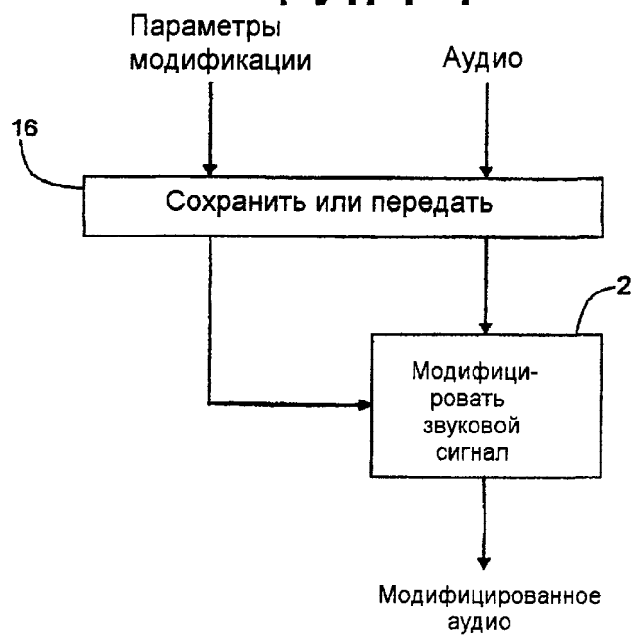
ФИГ. 2



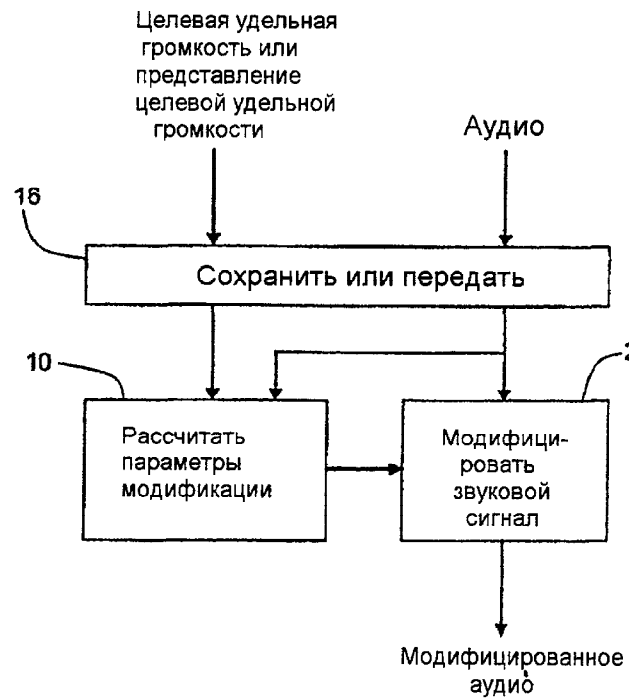
ФИГ. 3



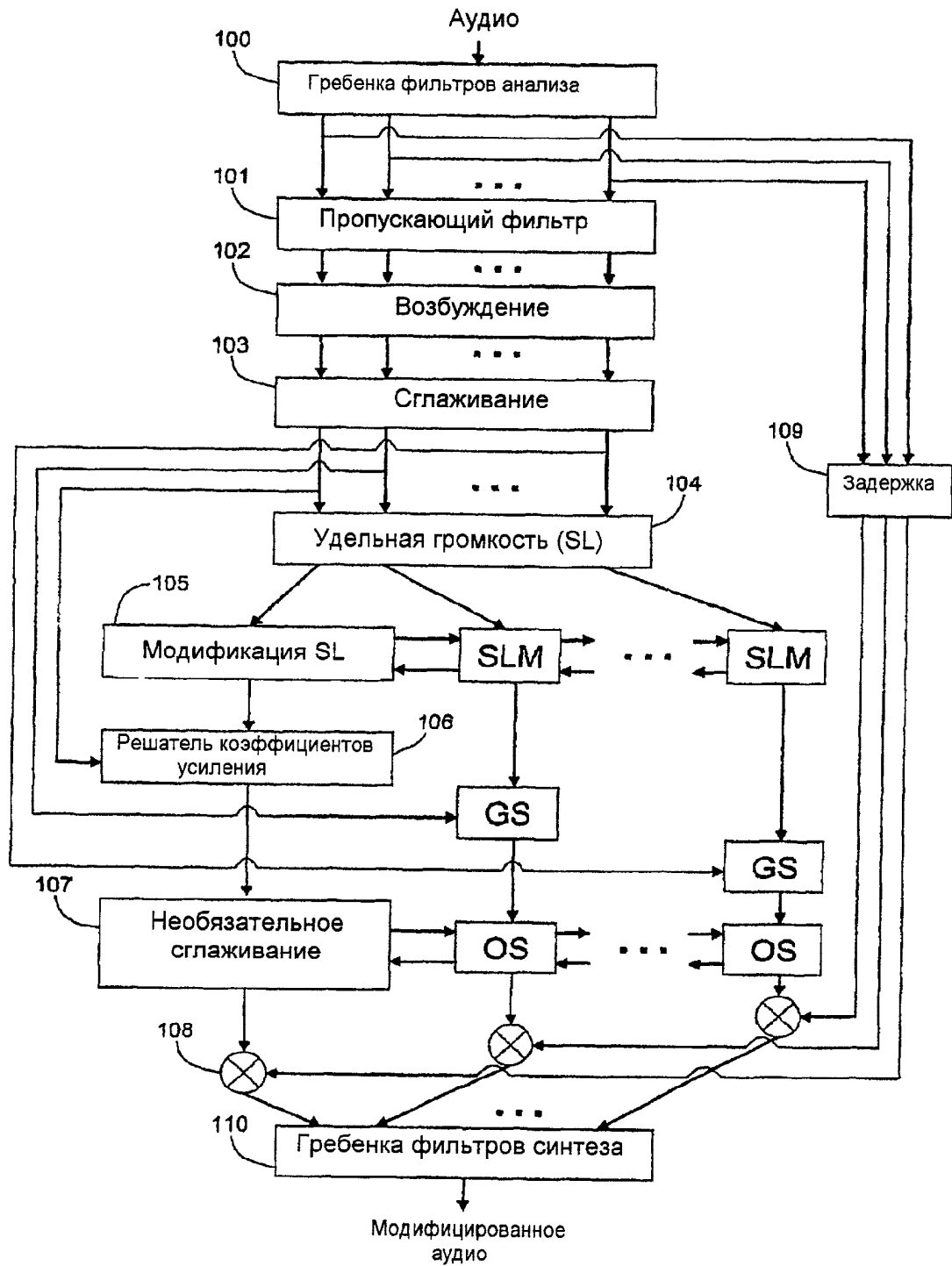
ФИГ. 4



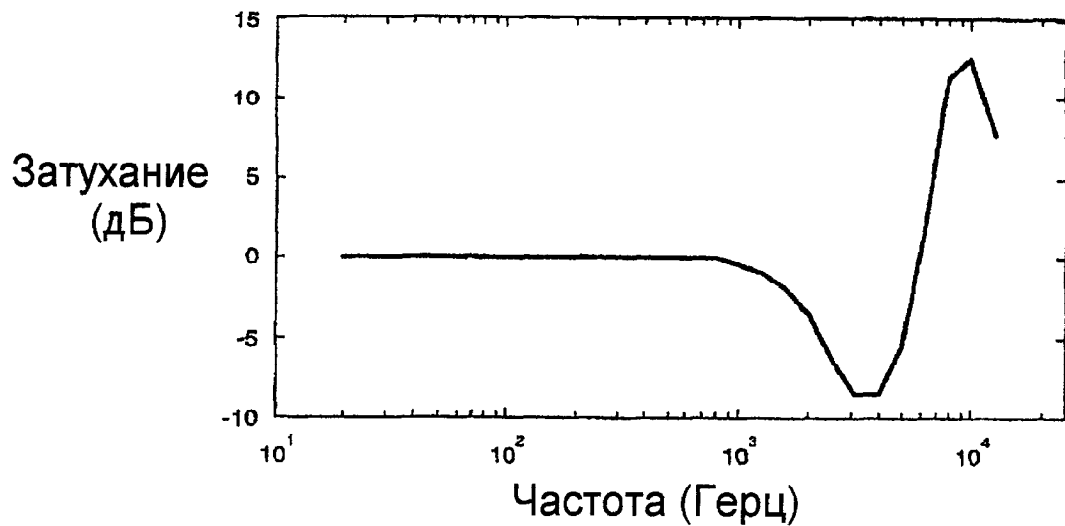
ФИГ. 5



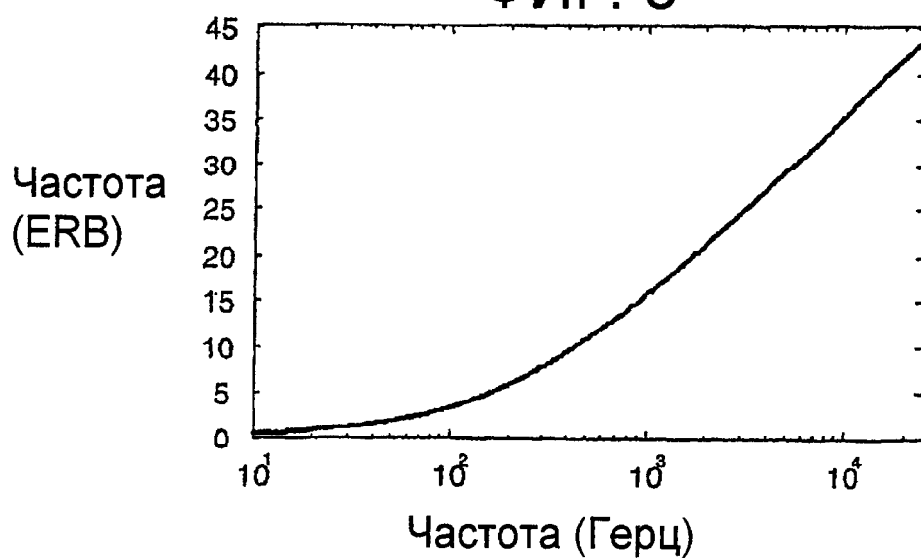
ФИГ. 6



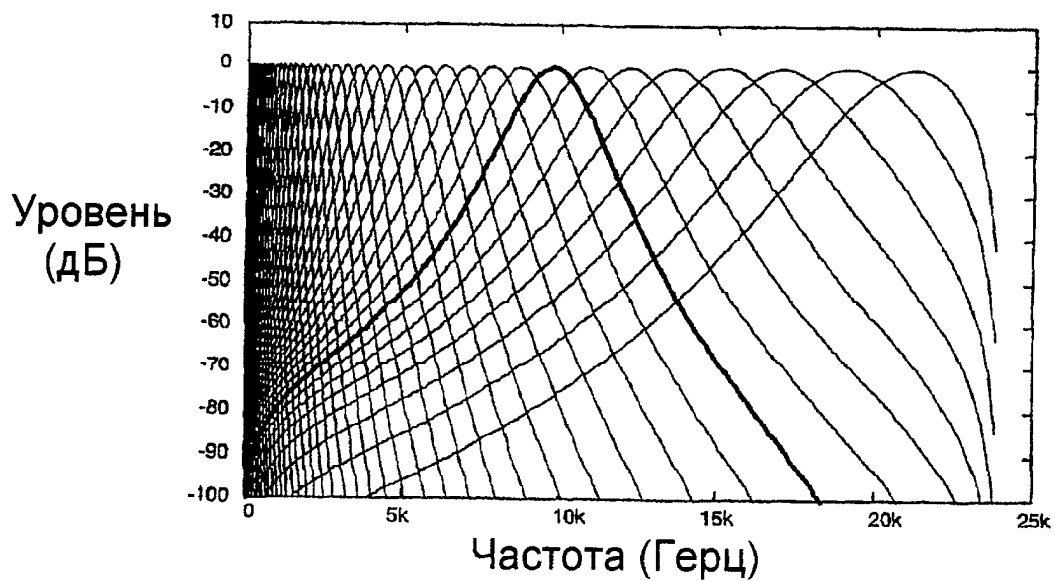
ФИГ. 7



ФИГ. 8

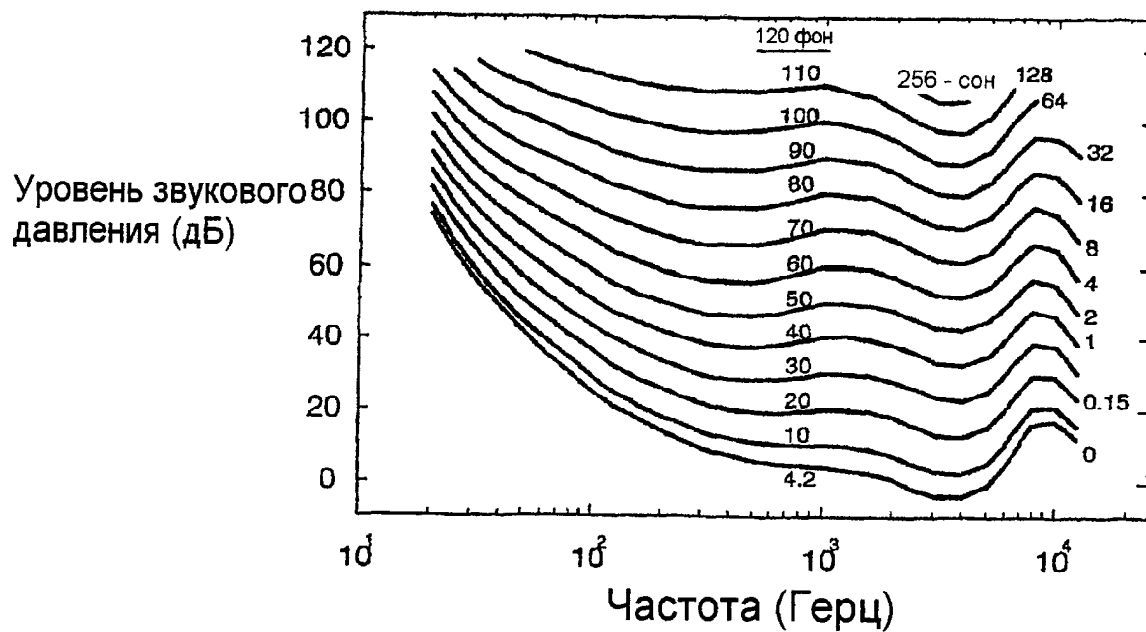


ФИГ. 9



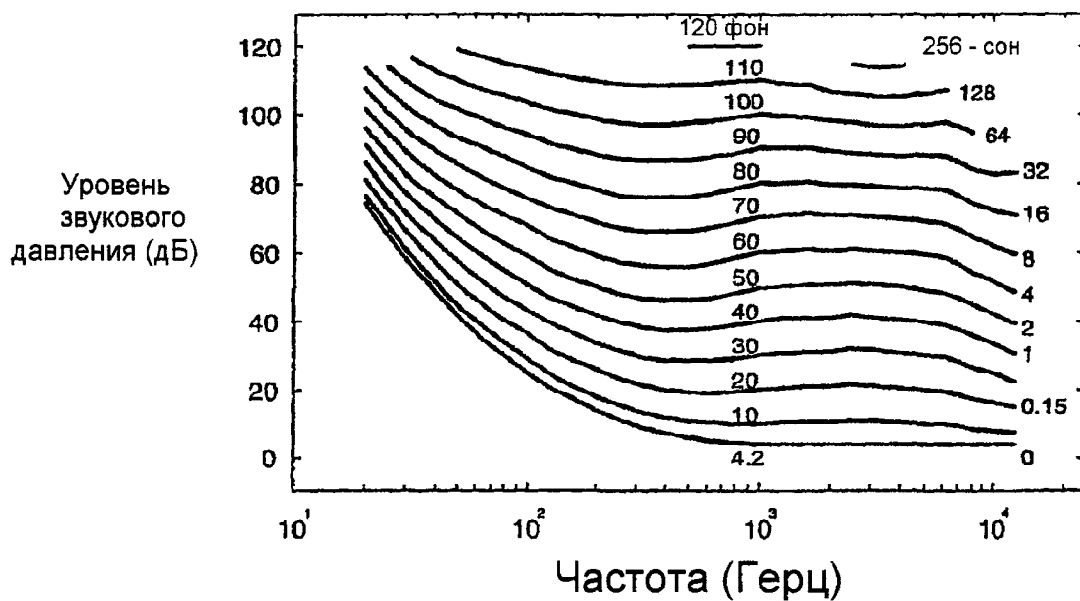
ФИГ. 10

«Кривые равной громкости» Приложение А ISO226

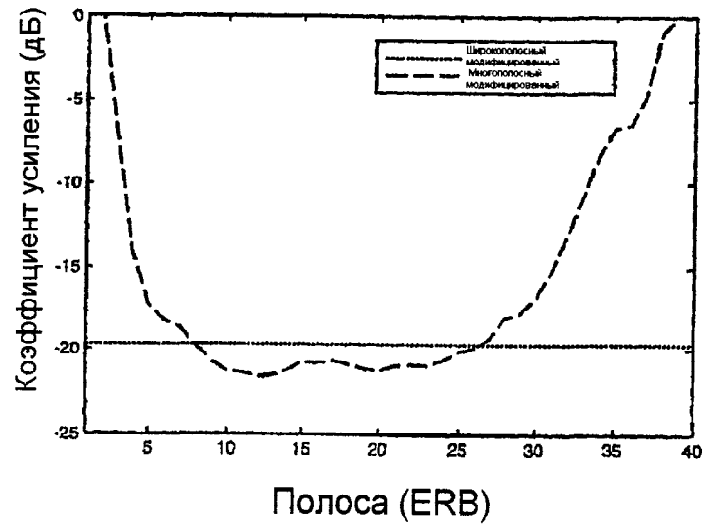


ФИГ. 11

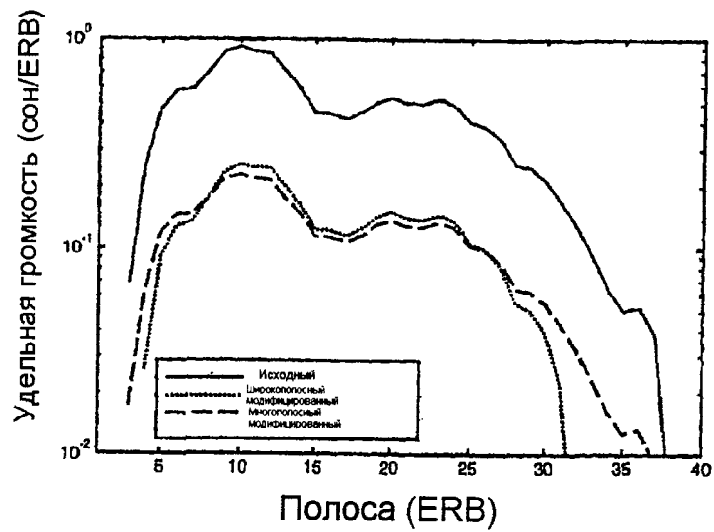
«Кривые равной громкости» Приложение А ISO226
(нормализованные пропускающим фильтром)



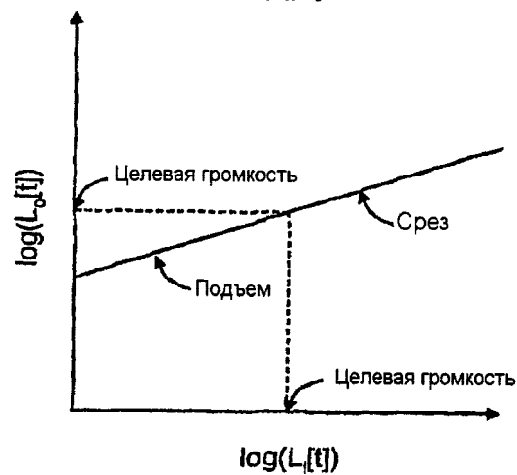
ФИГ. 12



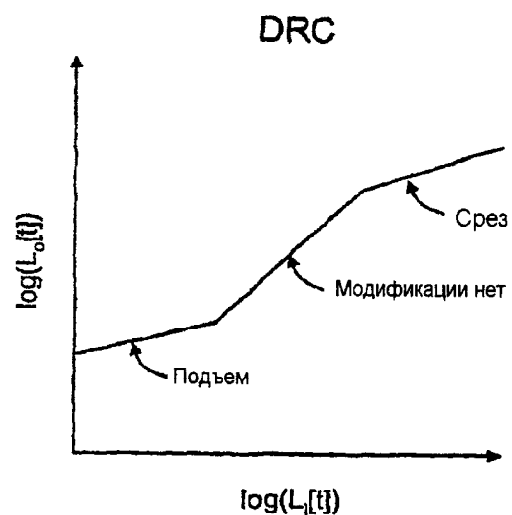
ФИГ. 13а



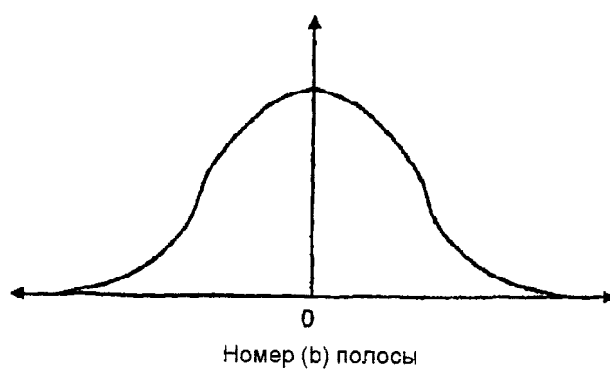
ФИГ. 13b
АРУ



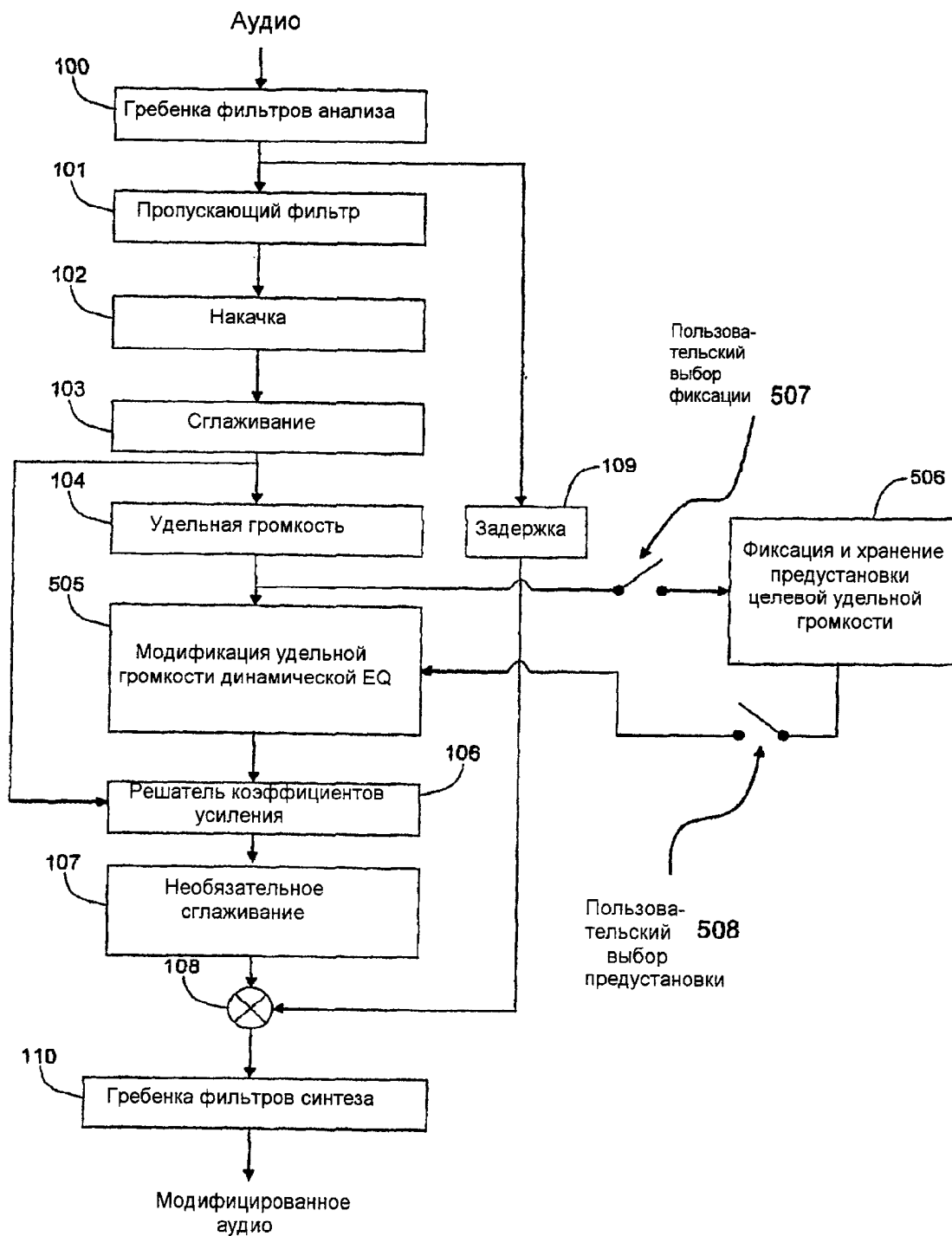
ФИГ. 14а



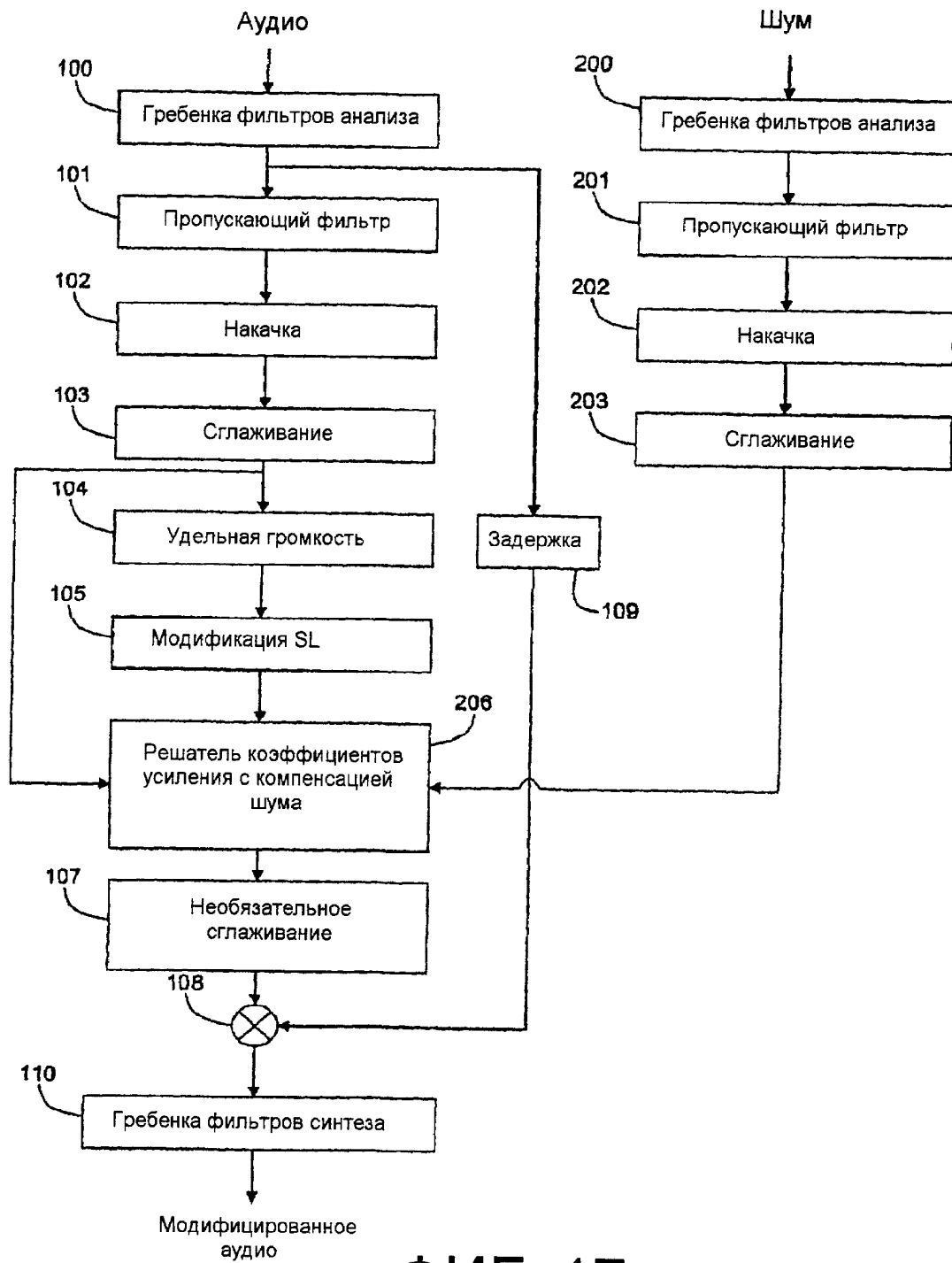
ФИГ. 14b
 $Q(b)$



ФИГ. 15



ФИГ. 16



ФИГ. 17