



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

**(12) ОПИСАНИЕ ИЗОБРЕТЕНИЯ К ПАТЕНТУ**

(21)(22) Заявка: **2009118955/09, 25.09.2007**

(24) Дата начала отсчета срока действия патента:  
**25.09.2007**

Приоритет(ы):

(30) Конвенционный приоритет:  
**20.10.2006 US 60/853,127**

(45) Опубликовано: **27.02.2011** Бюл. № 6

(56) Список документов, цитированных в отчете о поиске: **WO 2005/104360 A1, 03.11.2005. RU 2156538 C2, 20.09.2000. US 5369711 A, 29.11.1994. US 6700982 B1, 02.03.2004. US 5903655 A, 11.05.1999. US 5422602 A, 06.06.1995. GB 2215927 A, 27.09.1989. EP 0972340 A1, 19.01.2000.**

(85) Дата начала рассмотрения заявки РСТ на национальной фазе: **20.05.2009**

(86) Заявка РСТ:  
**US 2007/020747 (25.09.2007)**

(87) Публикация заявки РСТ:  
**WO 2008/051347 (02.05.2008)**

Адрес для переписки:  
**129090, Москва, ул.Б.Спасская, 25, стр.3,  
ООО "Юридическая фирма Городиский и  
Партнеры", пат.пов. А.В.Мишу, рег.№ 364**

(72) Автор(ы):

**СИФЕЛЬДТ Алан Джеффри (US),  
ГАНДРИ Кеннет Джеймс (US)**

(73) Патентообладатель(и):

**ДОЛБИ ЛЭБОРЕТЕРИЗ ЛАЙСЕНСИНГ  
КОРПОРЕЙШН (US)**

**(54) ОБРАБОТКА ДИНАМИЧЕСКИХ СВОЙСТВ АУДИО С ИСПОЛЬЗОВАНИЕМ  
ПЕРЕНАСТРОЙКИ**

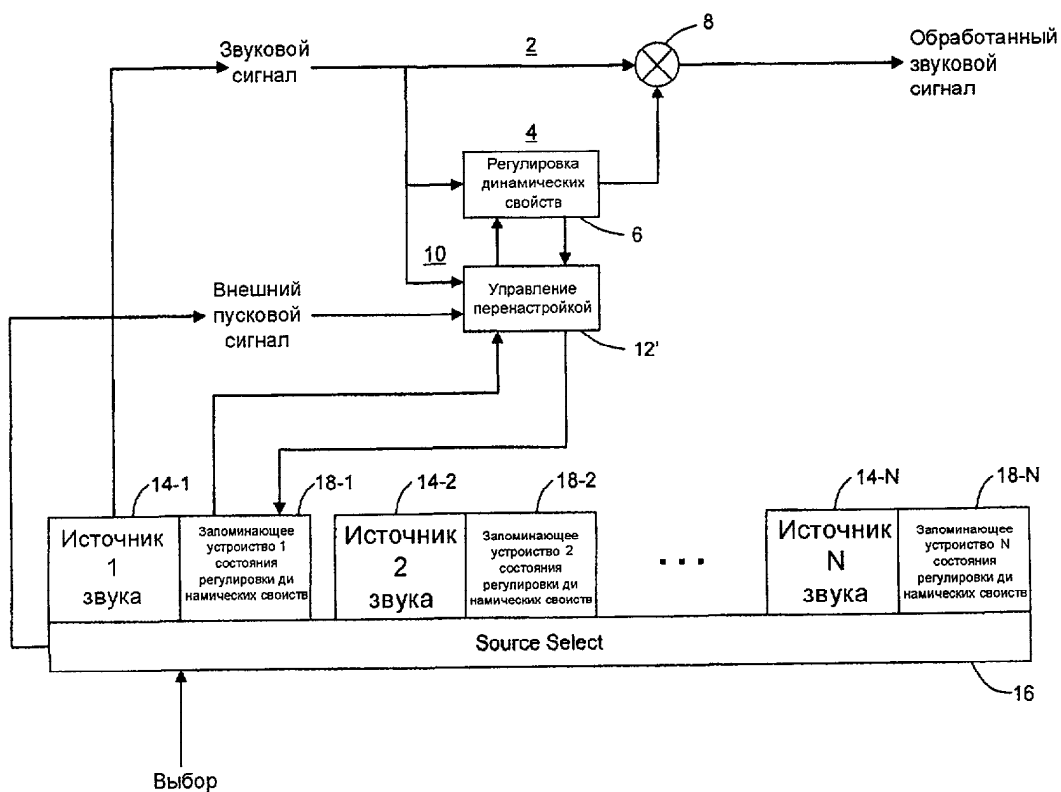
(57) Реферат:

Изобретение относится к обработке звуковых сигналов и предназначено для перенастройки при обработке динамических свойств аудио для быстрой адаптации к изменениям контента в звуковом сигнале. Технический результат - обеспечение адаптации к изменениям контента в звуковом сигнале. В процессоре или способе обработки динамических свойств аудио используют механизм или последовательность операций

перенастройки, для того чтобы быстро адаптироваться к изменениям контента в звуковом сигнале. Сигнал перенастройки может формироваться посредством анализа самого звукового сигнала, или перенастройка может запускаться от внешнего события, такого как изменение канала на телевизионном приемнике или изменение выбора входа на аудио/видеоприемнике. В случае внешнего пускового сигнала один или более признаков состояния процессора динамических свойств

для текущего источника звука могут сохраняться и ассоциативно связываться с таким источником звука перед переключением на новый источник звука. Затем, если система переключается обратно на первый источник

звuka, процессор динамических свойств может быть перенастроен на состояние, сохраненное ранее, или его приближение. 2 н. и 12 з.п.ф-лы, 6 ил.



ФИГ.3



FEDERAL SERVICE  
FOR INTELLECTUAL PROPERTY,  
PATENTS AND TRADEMARKS

**(12) ABSTRACT OF INVENTION**(21)(22) Application: **2009118955/09, 25.09.2007**(24) Effective date for property rights:  
**25.09.2007**

Priority:

(30) Priority:  
**20.10.2006 US 60/853,127**(45) Date of publication: **27.02.2011 Bull. 6**(85) Commencement of national phase: **20.05.2009**(86) PCT application:  
**US 2007/020747 (25.09.2007)**(87) PCT publication:  
**WO 2008/051347 (02.05.2008)**

Mail address:

**129090, Moskva, ul.B.Spasskaja, 25, str.3, OOO  
"Juridicheskaja firma Gorodisskij i Partnery",  
pat.pov. A.V.Mitsu, reg.№ 364**

(72) Inventor(s):

**SIFEL'DT Alan Dzheffri (US),  
GANDRI Kennet Dzhejms (US)**

(73) Proprietor(s):

**DOLBI LEhBORETERIZ LAJSENSING  
KORPOREJShN (US)**

**(54) PROCESSING DYNAMIC PROPERTIES OF AUDIO USING RETUNING**

(57) Abstract:

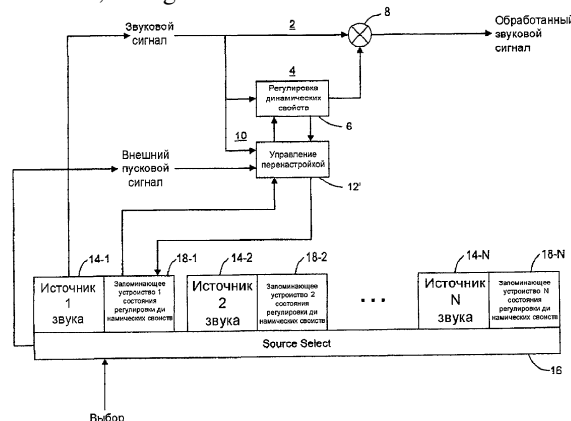
FIELD: physics.

SUBSTANCE: processor or method of processing dynamic properties of audio employs a mechanism or sequence of retuning operations for fast adaptation to changes in content in the audio signal. The retuning signal can be generated by analysing the audio signal itself, or retuning can be performed based on an external event, such as changing the channel on a television receiver or changing the input selection on an audio/video receiver. If there is an external activating signal, one or more features of the state of the dynamic property processor for the current audio source can be stored or associatively linked to such an audio source before switching to a new audio source. Further, if the system is switched back to the first audio source, the dynamic property processor can be

retuned to a state stored earlier or approximation thereof.

EFFECT: adaptation to changes in content in an audio signal.

14 cl, 6 dwg



ФИГ.3

## ОБЛАСТЬ ТЕХНИКИ, К КОТОРОЙ ОТНОСИТСЯ ИЗОБРЕТЕНИЕ

Изобретение относится к обработке звуковых сигналов. В частности, изобретение относится к процессору или способу обработки динамических свойств аудио, который использует механизм или последовательность операций перенастройки, для того чтобы быстро адаптироваться к изменениям контента в звуковом сигнале. Сигнал перенастройки может формироваться посредством анализа самого звукового сигнала, или перенастройка может запускаться от внешнего события, такого как изменение канала на телевизионном приемнике или изменение выбора входа на аудио/видеоприемнике. В случае внешнего пускового сигнала один или более признаков состояния процессора динамических свойств для текущего источника звука могут сохраняться и ассоциативно связываться с таким источником звука перед переключением на новый источник звука. Затем, если система переключается обратно на первый источник звука, процессор динамических свойств может быть перенастроен на состояние, сохраненное ранее, или его приближение. Изобретение также относится к компьютерным программам для осуществления таких способов на практике или управления таким устройством.

## УРОВЕНЬ ТЕХНИКИ

Цель обработки динамических свойств аудио состоит в том, чтобы менять уровень или динамические свойства звукового сигнала, чтобы были в пределах некоторых желательных пределов. Это в целом достигается созданием зависящего от времени измерения уровня звукового сигнала (например, среднеквадратического уровня или пикового уровня), а затем вычислением и применением зависящей от времени модификации сигнала (например, изменения коэффициента усиления), которая является функцией оценки уровня. Процессоры динамических свойств, применяющие такой режим работы, включают в себя автоматическую регулировку усиления (APU, AGC), регулировку динамического диапазона (УДД, DRC), расширители, ограничители, шумовые заслоны и т.д. Различные типы обработки динамических свойств сигналов изложены в заявке PCT/US 2005/038579 на выдачу международного патента Алана Джеффри Шифельда, опубликованной в качестве WO 2006/047600 4 мая 2006 года. Заявка указывает Соединенные Штаты среди других субъектов права. Заявка настоящим включена в состав посредством ссылки во всей своей полноте.

Фиг.1 изображает высокоуровневую структурную схему типового процессора динамических свойств аудио. Может считаться, что процессор должен иметь два тракта, верхний «сигнальный» тракт 2 и нижний тракт 4 «управления». В нижнем тракте последовательность операций управления или контроллер 6 динамических свойств («Элемент регулировки динамических свойств») измеряет уровень звукового сигнала и формирует один или более зависящих от времени параметров модификации в качестве функции измерения уровня. Как показано, параметры модификации выводятся из входного звукового сигнала. В качестве альтернативы параметры модификации могут выводиться из обработанного (выходного) аудио или из комбинации входного и выходного звуковых сигналов. В верхнем звуковом тракте 2 параметры модификации, сформированные элементом 6 регулировки динамических свойств, применяются к аудио для формирования обработанного аудио. Применение параметров модификации к звуковому сигналу может выполняться многими известными способами и, в общем, показано символом 8 умножителя. Например, в случае устройства или последовательности операций автоматической регулировки усиления может быть единственный параметр модификации широкополосного коэффициента усиления, который регулирует коэффициент усиления устройства или

последовательности операций переменного усиления/ослабления в основном тракте. На практике аудио также может задерживаться перед применением параметров модификации, для того чтобы компенсировать любую задержку, ассоциативно связанную с вычислением параметров модификации в последовательности операций 5 регулировки динамических свойств. Для простоты в представлении задержка не показана на фиг.1 или других фигурах в материалах настоящей заявки.

В последовательности операций регулировки динамических свойств типично, что как измерение уровня сигнала, так и результирующие параметры модификации, 10 вычисляются непрерывно во времени. В дополнение любое или оба измерения уровня сигнала и параметров модификации обычно сглаживаются по времени, чтобы минимизировать воспринимаемые артефакты от внесения в обработанный звук. Сглаживание, чаще всего, выполняется с использованием «быстрого нарастания» и «медленного спада», означающих, что параметры модификации изменяются 15 относительно быстро в ответ на увеличение уровня сигнала и медленно реагируют, в то время как уровень сигнала уменьшается. Такое сглаживание происходит в соответствии с динамическими свойствами естественных звуков и способом, которым люди воспринимают изменения громкости во времени. Следовательно, такое 20 временное сглаживание является почти универсальным в процессорах динамических свойств аудио.

Для некоторых приложений обработки динамических свойств постоянные времени, ассоциативно связанные с таким сглаживанием, могут быть довольно большими, 25 порядка одной или более секунд. Например, АРУ может вычислять оценку долговременного среднего уровня сигнала с использованием больших постоянных времени, а затем использовать результирующую оценку для формирования медленно меняющихся параметров модификации, которые перемещают средний уровень звука ближе к требуемому целевому уровню. В этом случае большие постоянные времени 30 могут быть желательны для того, чтобы сохранять кратковременные динамические свойства звукового сигнала. Предположим, что такое АРУ оперирует со звуком телевизионного приемника с целью поддержания согласующегося среднего уровня по программам и по различным каналам. В такой ситуации контент звукового сигнала, обрабатываемого АРУ, может внезапно изменяться или содержать скачок, например, 35 когда меняется канал, а потому ассоциативно связанный средний уровень звукового сигнала также может внезапно изменяться или содержать скачок. Со своими большими постоянными времени, однако, АРУ требует значительного количества времени для схождения к новому уровню и приведения модифицированного уровня 40 обработанного звука в соответствие с требуемым целевым уровнем. В течение такого времени адаптации зритель телевидения может воспринимать уровень звука слишком громким или слишком приглушенным. Как результат, зритель может быстро доставать пульт дистанционного управления для настройки только уровня громкости, чтобы оказаться борющимся с АРУ, в то время как она стремится к пределу.

Типичное решение предшествующего уровня техники для только что описанной 45 проблемы включает в себя использование постоянных времени, которые приспособливаются на основании динамических свойств сигнала. Например, если кратковременный уровень сигнала является значительно большим или меньшим, чем сглаженный уровень, как определено несколькими границами пороговых значений 50 вокруг сглаженного уровня, то операция сглаживания переключается на более быстрые постоянные времени нарастания и/или спада соответственно до тех пор, пока кратковременный уровень не отступает в пределы границ пороговых значений вокруг

сглаженного уровня. Следовательно, система переключается обратно на исходные, более медленные постоянные времени. Такая система может сокращать время адаптации АРУ, но пороговые значения и более короткие постоянные времени должны тщательно выбираться. Вообще, для любых приемлемых пороговых значений могут существовать сигналы, в которых исходные требуемые динамические свойства сигнала отклоняются за пределы границ пороговых значений вокруг среднего уровня, таким образом заставляя последовательность операций сглаживания ложно переключаться в режим быстрого нарастания или спада. Вследствие возможно частого возникновения такого ложного переключения постоянные времени режима быстрого нарастания и спада не должны выбираться слишком короткими, для того чтобы избежать нестабильности АРУ во время нормального программного материала. Как результат, сходжение АРУ во время внезапных переходных процессов или скачков в звуковом контенте все же может не быть настолько быстрым, насколько требуется.

Поэтому цель настоящего изобретения состоит в том, чтобы предложить лучшее решение для проблемы времени адаптации обработки динамических свойств в течение изменений звукового контента.

## РАСКРЫТИЕ ИЗОБРЕТЕНИЯ

Согласно аспекту изобретения способ обработки звукового сигнала (или устройство, осуществляющее способ на практике) включает в себя изменение динамических свойств звукового сигнала в соответствии с последовательностью операций регулировки динамических свойств, обнаружение изменения в контенте звукового сигнала и перенастройку последовательности операций регулировки динамических свойств в ответ на обнаружение такого изменения. Обнаружение изменения в контенте звукового сигнала может содержать одно или оба из (1) обнаружения события во временном развитии звукового сигнала, при котором уровень звукового сигнала остается ниже порогового значения,  $L_{\text{silence}}$ , на всем протяжении временного интервала, не меньшего чем первое пороговое значение времени,  $t_{\text{silence}}$ , и (2) обнаружения события во временном развитии звукового сигнала, при котором уровень звукового сигнала снижается на величину, большую чем порог различимости,  $L_{\text{drop}}$ , в пределах временного интервала, не большего чем второе пороговое значение времени,  $t_{\text{drop}}$ .

Согласно еще одному аспекту изобретения способ обработки звукового сигнала (или устройство, осуществляющее способ на практике) включает в себя изменение динамических свойств звукового сигнала в соответствии с последовательностью операций регулировки динамических свойств, прием подаваемого снаружи сигнала запуска перенастройки, указывающего изменение в звуковом сигнале, и перенастройку последовательности операций регулировки динамических свойств в ответ на сигнал запуска перенастройки. Сигнал запуска перенастройки может указывать одно или более из (1) действия пользователя, изменяющего канал, (2) действия пользователя, изменяющего источники входного сигнала, (3) действия пользователя, выбирающего функцию воспроизведения, перемотки назад или перемотки вперед, (4) переключения с одного файла на другой, (5) изменения программы, (6) переключения с одного формата кодирования звука на другой и (7) изменения параметров кодирования.

Согласно любому из вышеприведенных аспектов перенастройка последовательности операций регулировки динамических свойств может убыстрять скорость, с которой обработка динамических свойств адаптируется к изменениям во

входном сигнале. Перенастройка последовательности операций регулировки динамических свойств устанавливает значение одного или более параметров управления последовательностью операций или сигнал, функцией которого являются такие один или более параметров управления последовательностью операций, в хранимое значение или значение по умолчанию. Перенастройка последовательности операций регулировки динамических свойств может уменьшать значение одной или более постоянных времени, применяемых при определении одного или более параметров управления последовательностью операций, или сигнала, функцией которого являются такие один или более параметров управления последовательностью операций.

Хотя подход с адаптивными постоянными времени, описанный выше, может быть до некоторой степени полезным в сокращении времени адаптации процессора или последовательности операций динамических свойств аудио и может использоваться наряду с аспектами настоящего изобретения, такие процессор или последовательность операций могут быть значительно улучшены их наращиванием явными механизмом или последовательностью операций перенастройки, которые приводятся в действие при возникновении некоторых внезапных изменений или скачков в контенте звукового сигнала, являющегося обрабатываемым. Фиг.2 изображает настоящее изобретение в обобщенном виде, в котором дополнительный тракт управления, тракт 10 управления перенастройкой, добавлен в исходный процессор динамических свойств по фиг.1. Механизм или последовательность 12 операций управления перенастройкой («Элемент управления перенастройкой») может реагировать на один или оба из самого звукового сигнала (входного звукового сигнала или комбинации входного и выходного звуковых сигналов) и внешнего сигнала запуска. Таким образом, есть три возможные конфигурации: (1) сам звуковой сигнал (который, в свою очередь, может содержать входной звуковой сигнал или комбинацию входного и выходного звуковых сигналов), (2) внешний сигнал запуска и (3) оба, сам звуковой сигнал и внешний сигнал запуска. Если элемент 12 управления перенастройкой, механизм или последовательность операций управления перенастройкой, отправляет сообщение перенастройки в элемент 6 регулировки динамических свойств, давая ему указание перенастроить определенные аспекты его состояния регулировки динамических свойств некоторым образом, который предоставляет обработке динамических свойств аудио возможность быстро адаптироваться к новому уровню поступающего звукового сигнала. Перенастройка некоторых аспектов состояния регулировки динамических свойств, например, может (1) устанавливать одну или более постоянных времени в меньшее значение на короткий период времени (тем самым увыстря скорость адаптации обработки звука) и/или (2) устанавливать сглаженное измерение уровня звукового сигнала либо одного или более параметров модификации в сохраненное значение либо в среднее значение или значение по умолчанию.

Анализ звукового сигнала для определения события внезапного изменения или скачка в контенте сигнала

Решение о перенастройке может инициироваться элементом 12 управления перенастройкой некоторым количеством способов, например, посредством анализа самого поступающего звукового сигнала или посредством приема внешнего сигнала запуска перенастройки. При анализе звукового сигнала элемент 12 управления перенастройкой пытается обнаружить условия в сигнале, указывающие на внезапное изменение или скачок в контенте. При обнаружении такого условия с определенной

степенью достоверности элемент 12 управления перенастройкой принимает решение о перенастройке и впоследствии дает элементу 6 регулировки динамических свойств указание перенастроить один или более аспектов его состояния регулировки динамических свойств. Хотя другие технологии могут использоваться для

5 обнаружения внезапных изменений или скачков в контенте сигнала, ниже описаны два практических и эффективных примера для выполнения этого.

Использование внешнего пускового сигнала для определения события внезапного изменения или скачка в контенте сигнала

10 Во многих применениях высоконадежная внешняя информация может быть имеющейся в распоряжении для запуска перенастройки. Например, в телевизионном приемнике или кабельной «телевизионной абонентской приставке» действие пользователя, изменяющего канал, может служить в качестве внешнего пускового сигнала перенастройки. К тому же в телевизионном или аудио/видеоприемнике

15 действие пользователя, изменяющего источники входного сигнала, например, с «Видео 1» на «Видео 2», может служить для запуска перенастройки. Другие примеры включают в себя основанный на файлах медиаплеер, например портативное звуковое устройство или цифровой видеомаягнитофон. В таком случае перенастройка может

20 запускаться извне, когда система завершает воспроизведение одного файла и переключается на другой. Перенастройка также могла бы запускаться, когда пользователь сам или сама явным образом изменяет файл, нажимает воспроизведение либо быструю перемотку вперед или перемотку назад в новое местоположение на участке контента. В случае, в котором звук, являющийся обрабатываемым, был

25 декодирован из сжатого потока цифровых данных, например звука в формате Dolby Digital, могут существовать другие источники внешнего сигнала запуска перенастройки. «Dolby» и «Dolby Digital» являются торговыми марками корпорации Dolby Laboratories Licensing. Например, система кодирования звука

30 формата Dolby Digital является обязательным стандартом записи звука для телевидения высокой четкости в Соединенных Штатах. На заданном канале программирования звук отправляется непрерывно, но формат Dolby Digital может изменяться на границах программы. Например, основная программа может кодироваться в формате канала 5.1, но коммерческая может кодироваться в

35 стереофоническом формате. Декодер Dolby Digital обязательно обнаруживает такие изменения и может пересылать информацию в процессор динамических свойств аудио по настоящему изобретению в качестве пускового сигнала перенастройки. Также могут применяться другие изменения в параметрах кодир, например изменения в

40 параметре нормализации диалога Dolby Digital, «DIALNORM», который обычно сохраняется постоянным для одиночного участка контента. Подобный внешний пусковой сигнал перенастройки может формироваться также из других форматов кодирования, таких как формат MP3 и AAC. В дополнение к только что

45 перечисленным источникам внешних пусковых сигналов перенастройки возможны другие, и изобретение не подразумевается ограниченным перечисленными таковыми.

В случае, когда изобретение используется в устройстве, в котором аудио может возникать из набора известных источников, изобретение может быть дополнительно

50 улучшено, как показано на фиг.3. По низу фигуры изображены  $N$  источников с 14-1 по 14- $N$  звука (с «Источника 1 звука» по «Источник  $N$  звука»), где предполагается, что один из  $N$  источников звука выбирается для воспроизведения через систему обработки динамических свойств посредством устройства или последовательности 16 операций выбора источника («Элемента выбора источника»). Такие источники звука могли бы

представлять, но не в качестве ограничения, различные каналы на телевизоре или телевизионной абонентской приставке, различные входы на аудио/видеоприемнике или различные файлы в портативном медиаплеере. С каждым источником звука хранятся признаки состояния одного или более аспектов последовательности операций регулировки динамических свойств, как они имели место в последний раз, когда воспроизводился конкретный источник звука. Такие аспекты состояния регулировки динамических свойств, например, могут включать в себя одно или более из (1) сглаженного измерения уровня звукового сигнала, (2) одного или более параметров модификации, являющихся результатом сглаженного измерения уровня звукового сигнала, и (3) одной или более из постоянных времени, применяемых при получении сглаженного измерения уровня звукового сигнала. В примерном варианте осуществления, описанном ниже, применяется сглаженное измерение уровня звукового сигнала, которое опосредованно оказывает воздействие на параметр модификации коэффициента усиления (в примере коэффициент усиления является функцией сглаженного уровня звукового сигнала). Когда другой источник звука выбирается благодаря последовательности операций выбора источника, перед фактическим переключением на новый источник звука, элемент 16 выбора источника отправляет внешний пусковой сигнал перенастройки на элемент 12' управления перенастройкой. Это, в свою очередь, побуждает элемент 12' управления перенастройкой фиксировать признаки текущего состояния последовательности операций регулировки динамических свойств для текущего воспроизведения источника звука. Элемент 12' управления перенастройкой затем сохраняет признаки состояния регулировки динамических свойств в N местах с 18-1 по 18-N хранения состояния регулировки динамических свойств, ассоциативно связанных с текущим выбором аудио (запоминающих устройствах с 1 по N состояния регулировки динамических свойств). Затем элемент 16 выбора источника переключается на новый источник звука и отправляет ассоциативно связанные хранимые признаки состояния регулировки динамических свойств в последовательность операций регулировки перенастройки, которые, в свою очередь, перенастраивают последовательность операций регулировки динамических свойств с использованием хранимых признаков состояния.

Посредством использования аспектов последнего состояния регулировки динамических свойств источника звука для перенастройки последовательности операций регулировки динамических свойств при переключении на такой источник могут достигаться лучшие эксплуатационные качества, чем в отсутствие применения такого начального состояния. Без такой информации последовательность операций регулировки динамических свойств полагается на поступающий звуковой сигнал после перенастройки, в то время как адаптируется к новому источнику звука. На кратковременной основе уровень звука может быстро флуктуировать, а потому параметры модификации также могут быстро флуктуировать по мере того, как последовательность операций регулировки динамических свойств пытается быстро адаптироваться к новому аудио. Посредством использования аспектов последнего состояния процессора динамических свойств, ассоциативно связанного с источником звука, для перенастройки процессора динамических свойств вероятно, что процессор динамических свойств будет начинать с состояния, которое ближе к состоянию, к которому он будет стремиться в конечном счете для такого источника звука. Как результат, быстрым постоянным времени адаптации после перенастройки не нужно уменьшаться настолько, насколько могло бы потребоваться в ином случае, тем

самым, с предоставлением улучшенной устойчивости без принесения в жертву общего времени адаптации. В качестве примера рассмотрим переключение между двумя телевизионными станциями, первой, со звуком на среднем уровне в -15 дБ относительно цифрового значения полной шкалы, и второй, со звуком на среднем уровне в -30 дБ, обе с диапазоном плюс или минус 5 дБ вокруг своих соответственных средних значений. Допустим, что как раз перед переключением на вторую станцию средний уровень первой станции находится на -13 дБ. Кроме того, допустим, что средний уровень второй станции находится на -33 дБ после переключения. Это перепад в -20 дБ. Если кому-то требуется, чтобы система обработки динамических свойств адаптировалась за половину секунды, например, скорость адаптации в -40 дБ в секунду требуется при отсутствии любой другой информации. Такая скорость довольно быстра и может привносить неустойчивость в обработанный звук. С другой стороны, предположим, что последнее состояние обработки динамических свойств, сохраненное у второй станции, соответствует уровню в -28 дБ. Затем, после перенастройки, последовательность операций регулировки динамических свойств может перенастраиваться этим состоянием и существует перепад всего лишь в -33 дБ - (-28 дБ) = -5 дБ. Таким образом, для приближения к требуемому уровню за половину секунды требуется скорость адаптации всего лишь в -10 дБ в секунду (требуя меньшего сокращения постоянной времени). Этот пример иллюстрирует преимущества в сохранении последнего состояния последовательности операций регулировки динамических свойств у источника звука, который давал в результате такое состояние.

Результат, когда запускается перенастройка

Запуск перенастройки убыстряет скорость адаптации обработки динамических свойств аудио (скорость, с которой обработка динамических свойств адаптируется к изменениям во входном сигнале). Например, когда запускается перенастройка, одна или более постоянных времени, ассоциативно связанных со скоростью адаптации, могут быстро переключаться на значительно меньшие значения, а затем плавно возвращаться к своим исходным большим значениям через заданный период времени. В качестве альтернативы одна или более постоянных времени могут переключаться на значительно меньшие значения и оставаться в таких значениях в течение заданного периода времени. В компоновке АРУ сигнал  $c[t]$  регулировки постоянной времени может моментально уменьшать постоянные времени, используемые при вычислении сглаженного среднего уровня. Например, постоянные времени сглаживания могут быть уменьшены в течение продолжительности приблизительно в одну секунду, следующую за перенастройкой. В примерном варианте осуществления сигнал  $c[t]$  регулировки постоянной времени может начинаться со значения в «1» (указывая воздействие максимального изменения на одну или более постоянных времени) при возникновении пускового сигнала перенастройки; когда он равен «1», коэффициенты возрастания и спада устанавливаются в значения, значительно меньшие, чем их номинальные значения. Так как  $c[t]$  затухает до нуля через короткий период времени, такой как одна секунда (длительность периода времени не является критической), коэффициенты видоизменяются обратно в свои номинальные обычные (неперенастроенные) значения. В качестве альтернативы или в дополнение при возникновении пускового сигнала перенастройки значение сглаженного среднего уровня, вычисленного последовательностью операций регулировки динамических свойств, может перенастраиваться на состояние, хранимое с конкретным источником звука, или на значение по умолчанию.

## **ОПИСАНИЕ ЧЕРТЕЖЕЙ**

Фиг.1 - схематическая функциональная структурная схема, показывающая, на высоком уровне, исходный процессор динамических свойств аудио.

Фиг.2 - схематическая функциональная структурная схема, показывающая в соответствии с аспектами настоящего изобретения пример варианта осуществления настоящего изобретения, в котором, в общем виде, дополнительный тракт управления, тракт управления перенастройкой, добавлен в исходный процессор динамических свойств по фиг.1.

Фиг.3 - схематическая функциональная структурная схема, показывающая в соответствии с аспектами настоящего изобретения пример варианта осуществления настоящего изобретения, в котором процессор динамических свойств аудио дополнительно улучшен.

Фиг.4 показывает примерную функцию входа/выхода для типичной автоматической регулировки усиления (АРУ).

Фиг.5 показывает набор частотных характеристик полос пропускания, выбранный для имитации критической полосовой фильтрации, наблюдаемой по базилярной мембране в человеческом ухе.

Фиг.6 - схематическая функциональная структурная схема, показывающая в соответствии с аспектами настоящего изобретения пример варианта осуществления АРУ по настоящему изобретению.

## **ОСУЩЕСТВЛЕНИЕ ИЗОБРЕТЕНИЯ**

Описанное изобретение применимо к широкому многообразию приложений обработки динамических свойств, таких как регулировка динамического диапазона, компрессоры, ограничители, расширители и т. д. Однако оно особенно полезно, когда применяется к АРУ, в котором типично большие постоянные времени имеют следствием медленную адаптацию в отсутствие настоящего изобретения. Таким образом, предпочтительные варианты осуществления описывают, в качестве примера, применение изобретения к АРУ.

### **Базовая реализация АРУ**

Базовая реализация АРУ вычисляет зависящую от времени оценку среднего уровня сигнала и вычисляет требуемый модифицированный выходной уровень, который является функцией входного уровня и требуемого целевого уровня. Зависящая от времени модификация сигнала затем может вычисляться в качестве функции перепада между входным уровнем и требуемым выходным уровнем. Функция, отображающая вход в выход, предназначена для приведения уровня модифицированного звука ближе к требуемому целевому уровню, а фиг.4 изображает такую функцию входа/выхода. Для входного сигнала выше целевого уровня АРУ требует ослабления сигнала, а для входного сигнала ниже целевого уровня АРУ требует усиления сигнала. Степень ослабления или усиления может управляться изменением наклона линии на фиг.4.

Теоретически, измерение, используемое для вычисления среднего уровня звука, должно коррелировать с человеческим восприятием громкости. Это может достигаться многими способами, например взвешенным измерением среднеквадратической мощности или измерением психоакустической громкости. Простое невзвешенное измерение среднеквадратической мощности является слегка менее точным, чем два только что упомянутых способа, но по-прежнему демонстрирует высокую степень корреляции с человеческим восприятием громкости для большинства звуковых сигналов реального мира. Вследствие своей вычислительной простоты невзвешенное измерение среднеквадратической мощности

используется в предпочтительном варианте осуществления, но не должно рассматриваться в качестве ограничивающего изобретение.

Хотя, в принципе, аспекты изобретения могут быть реализованы аналоговой и/или цифровой областях, насколько можно ожидать, практические реализации должны быть реализованы в цифровой области, в которой каждый из звуковых сигналов представлен отдельными отсчетами или отсчетами внутри блоков данных. Оценка входного уровня и соответствующего требуемого выходного уровня и параметры модификации сигнала могут вычисляться непрерывно для аналогового звукового сигнала либо на основе отсчет за отсчетом у цифрового сигнала, но для этого примерного варианта осуществления, взамен, желательно вычислять такие величины для следующих друг за другом перекрывающихся блоков отсчетов цифрового звукового сигнала. Это главным образом происходит вследствие того обстоятельства, что цифровая блочная обработка полезна для обнаружения условий перенастройки из самого сигнала, как описано позже. Давая возможность цифровому звуковому сигналу быть представленным посредством  $x[n]$ , перекрывающиеся блоки звукового сигнала могут быть вычислены в качестве

$$x[n, t] = w[n]x[n + tN/2] \quad \text{for} \quad 0 < n < N - 1$$

(1)

где  $N$  - длина блока,  $N/2$  - степень перекрытия между следующими друг за другом блоками,  $t$  - индекс блока, а  $w[n]$  - оконная функция, такая как синусное окно. Для сигналов, подвергаемых выборке на 44100 Гц, хорошо работает настройка  $N=512$  или  $N=1024$ . Дополнительные подробности о компоновке обработки цифрового звука с применением перекрывающихся блоков отсчетов изложены в патенте 5899969 США Филдера и других («Frame-based audio with gain-control words» («Основанное на блоках аудио со словами регулировки усиления»)), каковой патент настоящим включен в состав посредством ссылки во всей своей полноте. Как описано ниже, АРУ вычисляет зависящий от времени коэффициент  $G[t]$  усиления, который затем перемножается с каждым блоком сигнала. Каждый из этих модифицированных блоков затем прибавляется с перекрытием, чтобы создать заключительный модифицированный звуковой сигнал  $y[n + tN/2]$ :

$$y[n + tN/2] = G[t]x[n, t] + G[t-1]x[n + N/2, t-1] \quad \text{для} \quad 0 < n < N/2$$

(2)

В качестве первого этапа при вычислении коэффициента  $G[t]$  усиления изменяющийся во времени мгновенный уровень сигнала  $L[t]$  вычисляется в качестве среднеквадратической мощности каждого блока  $x[n, t]$ . В децибелах уровень вычисляется как

$$L[t] = 10 \log_{10} \left( \frac{1}{N} \sum_{n=0}^{N-1} x^2[n, t] \right) \quad (3)$$

где 0 дБ соответствует уровню цифровой прямоугольной волны полной шкалы.

Затем мгновенный уровень может сглаживаться с быстрым нарастанием и медленным спадом для формирования сглаженного среднего уровня звукового сигнала  $\bar{L}[t]$ :

$$\bar{L}[t] = \begin{cases} \alpha[t] \cdot \bar{L}[t-1] + (1 - \alpha[t])L[t] & L[t] \geq L_{\min} \\ \bar{L}[t-1] & L[t] < L_{\min} \end{cases}$$

(4a)

где

$$\alpha[t] = \begin{cases} \alpha_{attack} & L[t] \geq \bar{L}[t-1] \\ \alpha_{release} & L[t] < \bar{L}[t-1] \end{cases}$$

(4b)

$$\text{и} \\ \alpha_{release} > \alpha_{attack}$$

(4c)

Коэффициенты  $\alpha_{attack}$  и  $\alpha_{release}$  сглаживания выбираются, чтобы давать требуемые времена нарастания и спада. Одним из способов задать это является время половинного затухания сглаживающего фильтра, то есть время, которое отнимается, чтобы импульсная характеристика временного сглаживающего фильтра затухала до половины своего исходного значения. Хотя выбор  $\alpha_{attack}$ , соответствующей времени половинного затухания в 1 секунду, и  $\alpha_{release}$ , соответствующей времени половинного затухания в 4 секунды, хорошо работает для АРУ, значения не являются критическими. Значение  $L_{min}$  в уравнении 4а представляет минимальный уровень, выше которого должен быть мгновенный уровень  $L[t]$  сигнала, для того чтобы обновлялся сглаженный уровень. Это предохраняет сглаженный средний уровень от падения слишком низко, а соответствующий коэффициент усиления от возрастания слишком высоко, когда сигнал попадает в относительную тишину. При условии, что 0 dB представляет уровень цифровой прямоугольной волны полной шкалы, настройка  $L_{min} = -60$  dB является приемлемым выбором, хотя уровень не является критическим.

Из сглаженного среднего уровня  $\bar{L}[t]$  и функции  $F_{AGC}$  входа/выхода, как показано на фиг.4, вычисляется требуемый выходной уровень  $\bar{L}_{out}[t]$

$$\bar{L}_{out}[t] = F_{AGC} \{ \bar{L}[t] \}$$

(5)

В заключение коэффициент  $G[t]$  усиления вычисляется по перепаду между выходным уровнем  $\bar{L}_{out}[t]$  и входным уровнем  $\bar{L}[t]$ :

$$G[t] = 10^{(\bar{L}_{out}[t] - \bar{L}[t]) / 20}$$

(6)

Когда  $\bar{L}_{out}[t] > \bar{L}[t]$ , коэффициент усиления является большим, чем единица, означая, что сигнал усиливается, а когда  $\bar{L}_{out}[t] < \bar{L}[t]$ , коэффициент усиления является меньшим чем единица, означая, что сигнал ослабляется.

АРУ с перенастройкой

Как упомянуто выше, механизм или функция перенастройки может вызывать одно из двух или оба из двух действий: (1) установку одной или более постоянных времени в меньшее значение на короткий период времени (тем самым убыстрение скорости адаптации обработки звука) и (2) установку сглаженного измерения уровня звукового сигнала либо одного или более параметров модификации в сохраненное значение либо в среднее значение или значение по умолчанию, которое может быть выражено в качестве  $\bar{L}_{init}$  (см. уравнение 8а, приведенное ниже и описания перед и после такого уравнения). Прежде всего, описана перенастройка одной или более постоянных времени.

Так как параметр модификации, коэффициент  $G[t]$  усиления, является функцией сглаженного среднего уровня  $\bar{L}[t]$ , скорость, с которой описанная АРУ адаптируется к звуку, обусловлена временами нарастания и спада, используемыми при вычислении  $L[t]$ . Как установлено выше, такие постоянные времени сглаживания выбираются, чтобы быть относительно большими, порядка нескольких секунд, так что АРУ не

изменяет уровень звука слишком быстро в течение временного развития нормального звукового контента. Однако если звуковой контент изменяется внезапно или содержит скачок, АРУ может использовать слишком долгое время для адаптации, особенно если

новый контент находится на значительно ином уровне, чем предыдущий контент. Чтобы ускорить адаптацию, сигнал перенастройки, описанный ранее, может использоваться для запуска модификации одной или более постоянных времени, ассоциативно связанных со сглаживанием уровня. В этом случае, когда запускается перенастройка, одна или более постоянных времени, ассоциативно связанных со скоростью адаптации, могут быстро переключаться на значительно меньшие значения, а затем плавно возвращаться к своим исходным большим значениям (или их приближениям) через заданный период времени. Возможны другие способы для убыстрения скорости адаптации после того, как запущена перенастройка. Например, вместо плавного возврата к своим исходным значениям или их приближениям, постоянные времени могут сохраняться на своих меньших значениях через заданный период времени, а затем возвращаться непосредственно к своим исходным значениям или их приближениям. Еще одна возможность состоит в том, чтобы обеспечивать скачкообразный возврат постоянных времени через заданный период времени. Не подразумевается, что изобретение должно быть ограниченным способом, которым постоянные времени возвращаются к своим исходным значениям или их приближению.

Для запуска перенастройки может существовать двоичный сигнал  $R[t]$  перенастройки, такой  $R[t] = 0$  во время нормальной работы, и  $R[t] = 1$ , когда перенастройка требуется на временном блоке  $t$ . Перенастройка может запускаться согласно анализу звукового сигнала или внешним источником, как описано ранее. Подробности примеров запуска перенастройки посредством анализа сигнала описаны ниже. В случае внешнего пускового сигнала может существовать сохраненное состояние, которым может перенастраиваться последовательность операций динамических свойств, как показано на фиг.3. Для описанного предпочтительного варианта осуществления такое сохраненное состояние может состоять из значения сглаженного среднего уровня  $\bar{L}[t]$  в последнем временном блоке, в котором обрабатывался конкретный источник звука. Начальное состояние, извлеченное из запоминающего устройства во время запущенной извне перенастройки (или, в качестве альтернативы, начальное состояние может быть значением по умолчанию, применяется или нет запускаемая извне перенастройка), может быть представлено значением  $\bar{L}_{init}$ .

По двоичному сигналу  $R[t]$  перенастройки может формироваться сигнал регулировки постоянной времени и впоследствии использоваться для мгновенного сокращения постоянных времени, используемых при вычислении сглаженного среднего уровня  $\bar{L}[t]$ , тем самым убыстряя скорость адаптации обработки. Сигнал  $c[t]$  регулировки постоянной времени может вычисляться, так что он равен «1» в момент, когда происходит перенастройка ( $R[t] = 1$ ), а затем экспоненциально затухает до «0» со скоростью, управляемой коэффициентом затухания  $\lambda < 1$ :

$$c[t] = 1 \text{ if } R[t] = 1$$

(7a)

иначе,

$$c[t] = \begin{cases} \lambda c[t-1] & L[t] \geq L_{min} \\ c[t-1] & L[t] < L_{min} \end{cases}$$

(7b)

Коэффициент затухания  $\lambda$  может быть выбран, чтобы давать время половинного затухания, например, в одну секунду, каковое означает, что постоянные времени сглаживания уменьшены в течение продолжительности в приблизительно одну секунду, следующую за перенастройкой. Отметим также, что сигнал управления

может обновляться, только если уровень  $L[t] \geq L_{\min}$ , для того чтобы предохранять быструю адаптацию от появления во время относительной тишины.

Лучше чем использование постоянных коэффициентов нарастания и спада для сглаживания уровня сигнала, как показано в уравнении 4b, коэффициенты теперь могут меняться со временем в качестве функции сигнала  $c[t]$  регулировки постоянной времени. Когда  $c[t] = 1$ , означая, что только что произошла перенастройка, коэффициенты нарастания и спада могут устанавливаться равными значениям, значительно меньшим, чем значения  $\lambda_{\text{attack}}$  и  $\lambda_{\text{release}}$  по уравнению 4b (например, около десяти процентов исходных значений или менее). По мере того как  $c[t]$  затем затухает до «0», коэффициенты могут интерполироваться обратно в свои номинальные значения  $\lambda_{\text{attack}}$  и  $\lambda_{\text{release}}$ . По выбору, в дополнение, небольшие значения коэффициентов, используемые непосредственно после перенастройки, также могут меняться в зависимости от того, существует ли значение  $\bar{L}_{\text{init}}$  начального состояния. Если никакого сохраненного начального состояния не существует, то могут использоваться очень маленькие значения коэффициентов, а если сохраненное начальное состояние существует, то могут использоваться слегка большие значения коэффициентов. В последнем случае предполагается, что инициализация  $\bar{L}[t]$  значением  $\bar{L}_{\text{init}}$  начинается  $\bar{L}[t]$  со значения, более близкого к значению, к которому он сойдется в конечном счете. Следовательно, может использоваться слегка меньшее время адаптации, соответствующее большим коэффициентам, по сравнению со случаем, где не доступно никакое начальное состояние. Последовательность операций адаптивного сглаживания может быть выражена в качестве

$$\bar{L}[t] = \bar{L}_{\text{init}} \quad \text{if } R[t] = 1 \text{ и существует}$$

$$\bar{L}_{\text{init}} \quad (8a)$$

иначе,

$$\bar{L}[t] = \begin{cases} \gamma[t] \cdot \bar{L}[t-1] + (1 - \gamma[t])L[t] & L[t] \geq L_{\min} \\ \bar{L}[t-1] & L[t] < L_{\min} \end{cases}$$

(8b)

где

$$\gamma[t] = \begin{cases} \gamma_{\text{attack}}[t] & L[t] \geq \bar{L}[t-1] \\ \gamma_{\text{release}}[t] & L[t] < \bar{L}[t-1] \end{cases},$$

(8c)

$$\gamma_{\text{attack}}[t] = c[t]\beta_{\text{attack}}[t] + (1 - c[t])\alpha_{\text{attack}},$$

(8d)

$$\gamma_{\text{release}}[t] = c[t]\beta_{\text{release}}[t] + (1 - c[t])\alpha_{\text{release}},$$

(8e)

$$\beta_{\text{attack}}[t] = \begin{cases} \beta_{\text{afast}} & R[t] = 1 \text{ и существует } \bar{L}_{\text{init}} \\ \beta_{\text{aslow}} & R[t] = 1 \text{ не существует } \bar{L}_{\text{init}} \\ \beta_{\text{attack}}[t-1] & \text{иначе} \end{cases},$$

(8f)

$$\beta_{release}[t] = \begin{cases} \beta_{rfast} & R[t] = 1 \text{ и существует } \bar{L}_{init} \\ \beta_{rslow} & R[t] = 1 \text{ не существует } \bar{L}_{init} \\ \beta_{release}[t-1] & \text{иначе} \end{cases},$$

5 (8g)

$$\beta_{afast} < \beta_{aslow} \ll \alpha_{attack},$$

(8h)

и

$$10 \quad \beta_{rfast} < \beta_{rslow} \ll \alpha_{release}$$

(8i)

Значения  $\beta_{afast}$  и  $\beta_{rfast}$  представляют начальные коэффициенты сглаживания нарастания и спада непосредственно после перенастройки, когда не предусмотрено  
 15 никакого начального состояния,  $\bar{L}_{init}$ . Значения, соответствующие временам половинного затухания в 50 миллисекунд и 200 миллисекунд соответственно, были найдены пригодными к употреблению. Конкретные значения не являются критическими. Значения  $\beta_{aslow}$  и  $\beta_{rslow}$  представляют большие начальные  
 20 коэффициенты сглаживания нарастания и спада непосредственно после перенастройки, когда предусмотрено начальное состояние,  $\bar{L}_{init}$ . В этом случае были найдены пригодными к употреблению значения, соответствующие временам половинного затухания, равным удвоенным таковым у их более быстрых аналогов, 100 миллисекунд и 400 миллисекунд соответственно. В этом случае также конкретные  
 25 значения не являются критическими.

Обнаружение перенастройки посредством анализа сигнала

Наиболее надежный способ для запуска перенастройки в раскрытом изобретении состоит в том, чтобы принимать пусковой сигнал внешним образом, из механизма или  
 30 последовательности операций, которые известны непосредственно привязанными к изменению звукового контента. Несколько таких механизмов были обсуждены ранее, например изменение канала на телевизионном приемнике или изменение выбора входа на аудио/видеоприемнике. Во многих случаях, однако, никакие такие внешние  
 35 механизмы могут не быть в распоряжении, а потому система должна полагаться на анализ самого звукового сигнала, если должна запускаться перенастройка. Например, предположим, что раскрытое изобретение находится в телевизионном приемнике, к которому пользователь подключил внешнюю кабельную телевизионную абонентскую  
 40 приставку. Кабельная телевизионная абонентская приставка ответственна за подстройку и изменение каналов, и декодированный звук пересылается в телевизор только в качестве непрерывного аудиопотока. Телевизор поэтому не принимает никакой явной информации о том, когда выбран новый канал. Единственной имеющейся в распоряжении информацией является та, которая может логически выводиться из аудиопотока, который он принимает.

45 Относительно простой, но эффективный способ запуска перенастройки посредством анализа сигнала состоит в том, чтобы обнаруживать событие во временном развитии звукового сигнала, при котором уровень звукового сигнала остается ниже порогового значения,  $L_{silence}$ , на всем протяжении минимального  
 50 периода времени (временного интервала, не меньшего чем пороговое значение времени,  $t_{silence}$ ). Другими словами, для обнаружения периода относительно тишины, имеющего продолжительность по меньшей мере такой же длительности, как заданный пороговый временной интервал. Такое событие, насколько можно ожидать, должно

указывать внезапное изменение или скачок в контенте аудио. Многие устройства, в частности телевизионные абонентские приставки, приглушают звуковой сигнал на короткий период времени, когда изменяется источник звука. Действие приглушения снижает звуковой сигнал до уровня, гораздо меньшего, чем обычно встречается в нормальном звуковом контенте. В случае цифрового звука, например, оно может уменьшать отсчеты сигнала до нуля. Такие условия во временном развитии сигнала могут обнаруживаться посредством анализа кратковременного уровня  $L[t]$ , обсужденного ранее. Если  $L[t]$  остается ниже, чем пороговое значение  $L_{\text{silence}}$  на всем протяжении временного интервала в по меньшей мере  $t_{\text{silence}}$ , то может запускаться перенастройка. Этот подход может быть выражен как

$$R[t] = \begin{cases} 0 & s[t] < t_{\text{silence}} \\ 1 & s[t] \geq t_{\text{silence}} \end{cases}$$

(9a)

где сигнал  $s[t]$  вычисления паузы обновляется согласно

$$s[t] = \begin{cases} 0 & L[t] > L_{\text{silence}} \\ s[t-1] + 1 & L[t] \leq L_{\text{silence}} \end{cases}$$

(9b)

Если 0 дБ соответствует уровню цифровой прямоугольной волны полной шкалы, то настройка  $L_{\text{silence}}$ , например, равная -90 дБ, хорошо работает на практике, хотя уровень не является критическим. Настройка  $t_{\text{silence}}$  на время, например, в 0,25 секунд является приемлемым выбором для многих применений, хотя период времени не является критичным.

В некоторых ситуациях промежуток относительной тишины может не привноситься, когда звуковой контент внезапно изменяется или содержит скачок, а потому вышеприведенная технология может не быть эффективной при запуске перенастройки. Поэтому желательно альтернативное событие запуска перенастройки, чтобы иметь дело с дополнительными случаями. В любой ситуации, в которой звуковой контент внезапно изменяется или содержит разрыв, уровень нового аудио будет либо большим чем, равным либо меньшим чем уровень предыдущего звукового контента. Так как большинство процессоров динамических свойств применяют нарастание, которое значительно быстрее, чем его спад, случай, в котором уровень нового звукового контента является большим чем или равным старому, является менее проблематичным, чем когда новый звуковой контент является значительно более низким, чем старый. В последнем случае время медленного спада может иметь следствием неприемлемо длительное время адаптации. Поэтому запуск перенастройки, когда новый звуковой контент значительно ниже, чем старый, может считаться особенно желательным. Чтобы делать так, можно применять признак, общий для большинства звуковых сигналов, состоящих из естественных звуков. Естественные звуки могут проявлять внезапные повышения уровня, но они обычно демонстрируют более постепенное затухание. Звуки, которые внезапно прерываются, не являются самыми распространенными в большинстве звуковых записей. Однако если звуковой контент внезапно переключается, например, благодаря изменению выбора входа и уровень нового звукового входа является значительно меньшим, чем старый, то результирующий аудиопоток демонстрирует радикальное неестественное мгновенное падение уровня. Такое падение уровня может обнаруживаться по кратковременному уровню  $L[t]$  и использоваться для запуска перенастройки.

Таким образом, еще один эффективный способ запуска перенастройки посредством

анализа сигнала состоит в том, чтобы обнаруживать событие во временном развитии звукового сигнала, при котором уровень звукового сигнала снижается на величину, большую чем порог различимости,  $L_{drop}$ , в пределах временного интервала, не большего чем второе пороговое значение времени,  $t_{drop}$ . Более точно, если перепад между  $L[t]$  и  $L[t-1]$  является большим чем заданное пороговое значение,  $L_{drop}$ , в пределах временного интервала  $t_{drop}$ , то запускается перенастройка

$$R[t] = \begin{cases} 1 & L[t] - L[t-1] < L_{drop} \\ 0 & \text{иначе} \end{cases}$$

(10)

Подходящей разновременностью  $t_{drop}$  является время одного блока цифровой обработки. Например, с перекрытием блока в половину время одного блока соответствует  $N/2$  отсчетам. Если  $N=512$ , а частотой выборки является 48000 Гц,  $t_{drop}$  приблизительно равно 5 миллисекундам. Если  $N=1024$ ,  $t_{drop}$  равно приблизительно 10 миллисекундам. Перепад уровней,  $L_{drop}$ , в приблизительно от -10 дБ до -20 дБ был найден подходящим, хотя перепад уровней не является критичным.

Вышеприведенная технология эффективна при запуске перенастройки, когда уровень нового звукового контента значительно ниже, чем старого, но она иногда может быть чувствительной к ложному запуску перенастройки во время нормального звукового контента. Описанная технология ищет падение в суммарном уровне сигнала, и иногда доминирующая часть звукового сигнала, например удар в барабан, может демонстрировать такое поведение, в то время как оставшаяся часть сигнала не может. Доминирующая сигнальная составляющая, однако, осуществляет наиболее значительный вклад в общий уровень, а потому может заставлять запускаться перенастройку. Чтобы справляться с таким непредусмотренным запуском, схема обнаружения может быть улучшена обнаружением перенастройки, когда уровень сигнала демонстрирует значительное падение уровня одновременно по множеству полос частот (например, всем полосам частот) в пределах порогового периода времени. Этим способом все такие полосы частот осуществляют равный вклад в процесс обнаружения перенастройки, независимо от абсолютного уровня в любой данной полосе. Для реализации такой технологии сначала вычисляется мгновенный уровень в пределах множества полос частот. Этот уровень представлен в качестве  $L[b, t]$ , где  $b$  указывает индекс полосы. Он может эффективно формироваться вычислением дискретного преобразования Фурье (ДПФ, DFT) каждого блока  $x[n, t]$  аудио и перемножением ДПФ с частотной характеристикой  $C_h[k]$  полосы пропускания для каждой полосы  $b$ :

$$L[b, t] = 10 \log_{10} \left( \sum_{k=0}^{N-1} C_h^2[k] |X[k, t]|^2 \right)$$

(11a)

где  $X[k, t]$ , ДПФ от  $x[n, t]$ , вычисляется в качестве

$$X[k, t] = \sum_{n=0}^{N-1} x[n, t] e^{-j \frac{2\pi kn}{N}}$$

(11b)

Набор частотных характеристик  $C_h[k]$  полос пропускания преимущественно может выбираться для имитации критической полосовой фильтрации, наблюдаемой по базилярной мембране в человеческом ухе. Как предложено Муром и Гласбергом («A Model for the Prediction of Thresholds, Loudness, and Partial Loudness» («Модель для

предсказания пороговых значений, громкости и частичной громкости»), Brian Moore, Brian Glasberg, and Thomas Baer, J. Audio Eng. Soc, Vol. 45, No. 4, 1997 April), такие фильтры демонстрируют приблизительно округлую экспоненциальную форму и равномерно разнесены по шкале эквивалентных прямоугольных полос пропускания (ERB). Фиг.5 изображает пригодный набор фильтров с разнесением в 1 ERB, дающий в результате итоговую сумму в 40 полос.

По мгновенному уровню для каждой полосы,  $L[b, t]$ , может быть вычислен перепад между следующими друг за другом временными блоками, усредненными по всем полосам

$$D[t] = \frac{1}{B} \sum_{b=1}^B L[b, t] - L[b, t-1]$$

(12)

Для звуковых сигналов реального мира,  $D[t]$  является крайне малым, только если уровень сигнала значительно падает в большинстве полос  $b$ . Если перепад является меньшим, чем некоторое пороговое значение,  $D_{\text{drop}}$ , то обнаружена перенастройка

$$R[t] = \begin{cases} 1 & D[t] < D_{\text{drop}} \\ 0 & \text{иначе} \end{cases}$$

(13)

На практике хорошо работает настройка  $D_{\text{drop}}$  между -10 и -20 дБ, хотя настройка не является критической. Получающаяся в результате многополосная технология является менее чувствительной к ложному запуску перенастройки, чем более простая технология, которая проверяет только перепад в общем уровне сигнала.

Фиг.6 показывает структурную схему АРУ, воплощающей аспекты настоящего изобретения. Цифровой звуковой входной сигнал  $x[n, t]$  (см. фиг.1) подается в два тракта устройства или последовательности операций АРУ. В тракте управления устройство или последовательность 20 операций «Вычислить уровень» вычисляет измерение звука, которое может быть зависящим от времени мгновенным уровнем сигнала  $L[t]$ , вычисленным в качестве среднеквадратической мощности каждого блока входного сигнала  $x[n, t]$  (см. уравнение 3). Затем мгновенный уровень  $L[t]$  сглаживается по времени в устройстве или последовательности 22 операций «Сгладить уровень» для формирования сглаженного среднего уровня звукового сигнала  $\bar{L}[t]$  (см. уравнение 4а). В отсутствие перенастройки временное сглаживание может иметь характеристику быстрого нарастания и медленного спада. Сгладить уровень 22 может управляться входом «Сигнала  $c[t]$  регулировки постоянной времени» и входом «Инициализировать  $\bar{L}[t]$ ». Вход сигнала  $c[t]$  регулировки постоянной времени может заставлять одну или более из постоянных времени, применяемых в Сгладить уровень 22, изменяться при возникновении перенастройки, как описано выше. Вход инициализировать  $L[t]$  может заставлять выход у Сгладить уровень 22 принимать хранимое значение или значение по умолчанию  $\bar{L}_{\text{init}}$  при возникновении перенастройки, как описано выше. Выход  $\bar{L}[t]$  у Сгладить уровень 22 подается в устройство или функцию 24 «Применить функцию  $F_{\text{AGC}}$  входа/выхода» и в устройство или функцию 26 «Вычислить коэффициент усиления». Устройство или функция 24 применяет функцию, описанную в связи с фиг.4 для выдачи требуемого выходного уровня  $\bar{L}_{\text{out}}[t]$  (см. уравнение 5). Вычислить коэффициент усиления 26 вычисляет коэффициент  $G[t]$  усиления по перепаду между выходным уровнем  $\bar{L}_{\text{out}}[t]$  и входным уровнем  $\bar{L}[t]$  (см. уравнение 6). Когда  $\bar{L}_{\text{out}}[t] > \bar{L}[t]$ , коэффициент усиления является большим чем единица, означая, что сигнал усиливается, а когда  $\bar{L}_{\text{out}}[t] < \bar{L}[t]$ ,

коэффициент усиления является меньшим чем единица, означая, что сигнал ослабляется. Зависящий от времени коэффициент  $G[t]$  усиления, параметр модификации, затем подается в устройство или последовательность 28 операций «Умножить», где он затем перемножается с каждым блоком сигнала  $x[n, t]$  для создания модифицированного звукового сигнала  $y[n+tN/2]$  (см. уравнение 2). Хотя, как описано, параметр  $G[t]$  модификации коэффициента усиления является независимым от частоты, может вычислять зависящий от частоты параметр  $G[b, t]$  модификации коэффициента усиления, в котором  $b$  является индексом полосы. В этом случае

Вычислить уровень 20 может вычислять зависящий от частоты мгновенный уровень  $L[b, t]$  сигнала, Сгладить уровень 22 может выдавать зависящий от частоты выходной сигнал  $\bar{L}[b, t]$  (его входы управления Сигнал  $c[t]$  регулировки постоянной времени и Инициализировать  $L[t]$  также могут быть зависящими от частоты), Применить функцию  $F_{AGC}$  может применять зависящую от частоты функцию, а Вычислить коэффициент усиления 26 выдает зависящий от частоты изменяющийся во времени коэффициент  $G[b, t]$  усиления.

#### 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 состав 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 программах, 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 предопределенным 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 обработки звукового сигнала, состоящий в том, что:  
меняют динамические свойства звукового сигнала в соответствии с

последовательностью операций регулировки динамических свойств,  
обнаруживают первый тип события во временном развитии звукового сигнала, при котором уровень звукового сигнала снижается на величину, большую чем порог различимости,  $L_{drop}$ , в пределах временного интервала, не большего чем пороговое значение времени,  $t_{drop}$ , при этом упомянутое обнаружение обнаруживает снижение уровня звукового сигнала во множестве полос частот, и

перенастраивают последовательность операций регулировки динамических свойств в ответ на обнаружение такого первого типа изменения.

2. Способ по п.1, дополнительно содержащий этапы, на которых обнаруживают второй тип события во временном развитии звукового сигнала, при котором уровень звукового сигнала остается ниже порогового значения,  $L_{silence}$ , на всем протяжении временного интервала, не меньшего чем другое пороговое значение времени,  $t_{silence}$ , и перенастраивают последовательность операций регулировки динамических свойств в ответ на обнаружение такого второго типа изменения.

3. Способ по п.1, в котором перенастройка последовательности операций регулировки динамических свойств убыстряет скорость, с которой обработка динамических свойств адаптируется к изменениям во входном сигнале.

4. Способ по п.1, в котором перенастройка последовательности операций регулировки динамических свойств устанавливает значение одного или более параметров управления последовательностью операций или сигнал, функцией которого являются такие один или более параметров управления последовательностью операций, в хранимое значение или значение по умолчанию.

5. Способ по п.1, в котором перенастройка последовательности операций регулировки динамических свойств уменьшает значение одной или более постоянных времени, применяемых при определении одного или более параметров управления последовательностью операций, или сигнала, функцией которого являются такие один или более параметров управления последовательностью операций.

6. Способ по п.1, дополнительно содержащий этапы, на которых принимают подаваемый извне сигнал запуска перенастройки, указывающий изменение в звуковом сигнале, и перенастраивают последовательность операций регулировки динамических свойств в ответ на сигнал запуска перенастройки.

7. Способ по п.6, в котором сигнал запуска перенастройки указывает одно или более из  
действия пользователя, изменяющего канал,  
действия пользователя, изменяющего источники входного сигнала,

действия пользователя, выбирающего функцию воспроизведения, перемотки назад или перемотки вперед,

переключения с одного файла на другой,  
изменения программы,

переключения с одного формата кодирования звука на другой, и  
изменения параметров кодирования.

8. Устройство обработки звукового сигнала, содержащее  
средство изменения динамических свойств звукового сигнала в соответствии с

последовательностью операций регулировки динамических свойств,

средство обнаружения события во временном развитии звукового сигнала, при котором уровень звукового сигнала снижается на величину, большую чем порог различимости,  $L_{drop}$ , в пределах временного интервала, не большего чем пороговое значение времени,  $t_{drop}$ , при этом упомянутое средство обнаружения обнаруживает снижение уровня звукового сигнала во множестве полос частот, и

средство перенастройки последовательности операций регулировки динамических свойств в ответ на обнаружение такого изменения.

9. Устройство по п.8, дополнительно содержащее

средство обнаружения события во временном развитии звукового сигнала, при котором уровень звукового сигнала остается ниже порогового значения,  $L_{silence}$ , на всем протяжении временного интервала, не меньшего чем другое пороговое значение времени,  $t_{silence}$ , и

средство перенастройки последовательности операций регулировки динамических свойств в ответ на обнаружение такого изменения.

10. Устройство по п.8, в котором упомянутое средство перенастройки последовательности операций регулировки динамических свойств убыстряет скорость, с которой обработка динамических свойств адаптируется к изменениям во входном сигнале.

11. Устройство по п.8, в котором упомянутое средство для перенастройки последовательности операций регулировки динамических свойств устанавливает значение одного или более параметров управления последовательностью операций или сигнал, функцией которого являются такие один или более параметров управления последовательностью операций, в хранимое значение или значение по умолчанию.

12. Устройство по п.8, в котором упомянутое средство перенастройки последовательности операций регулировки динамических свойств уменьшает значение одной или более постоянных времени, применяемых при определении одного или более параметров управления последовательностью операций, или сигнала, функцией которого являются такие один или более параметров управления последовательностью операций.

13. Устройство по п.8, дополнительно содержащее  
средство приема подаваемого извне сигнала запуска перенастройки, указывающего изменение в звуковом сигнале, и

средство перенастройки последовательности операций регулировки динамических свойств в ответ на сигнал запуска перенастройки.

14. Устройство по п.13, в котором сигнал запуска перенастройки указывает одно или более из

действия пользователя, изменяющего канал, действия пользователя, изменяющего источники входного сигнала,

действия пользователя, выбирающего функцию воспроизведения, перемотки назад  
или перемотки вперед,  
переключения с одного файла на другой,  
изменения программы,  
5 переключения с одного формата кодирования звука на другой, и  
изменения параметров кодирования.

10

15

20

25

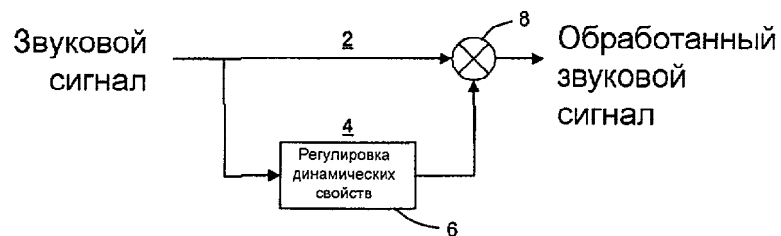
30

35

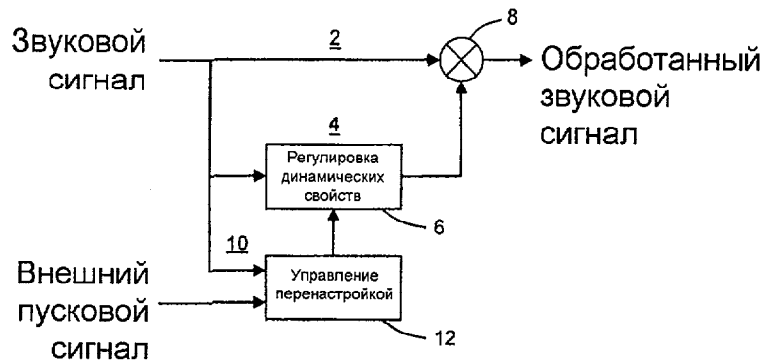
40

45

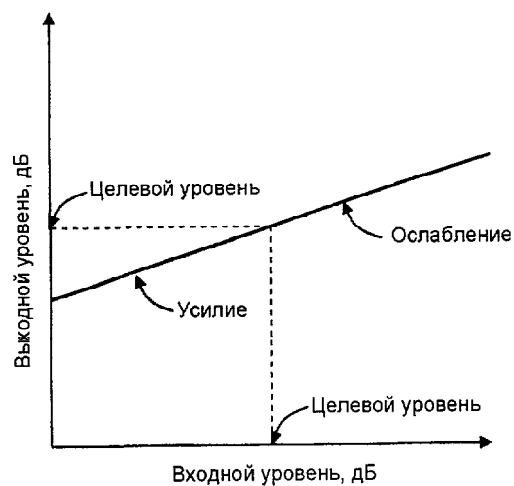
50



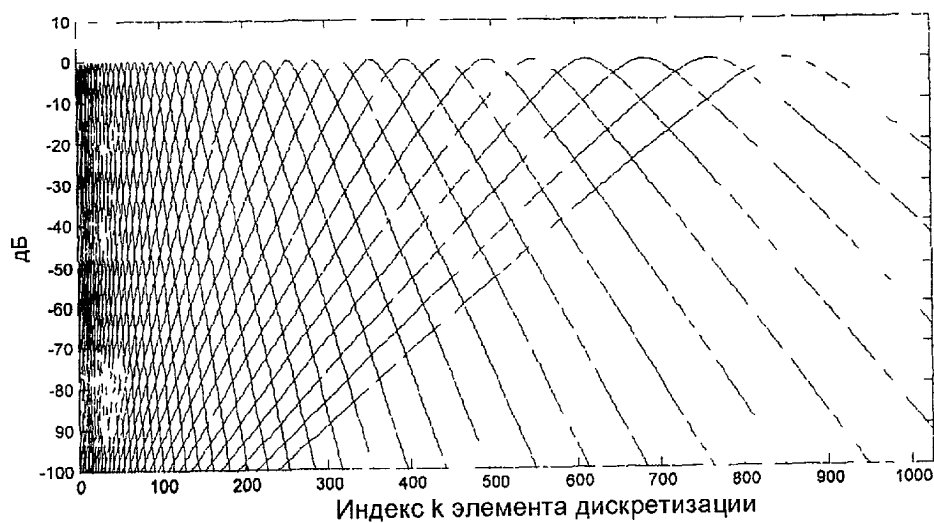
ФИГ.1



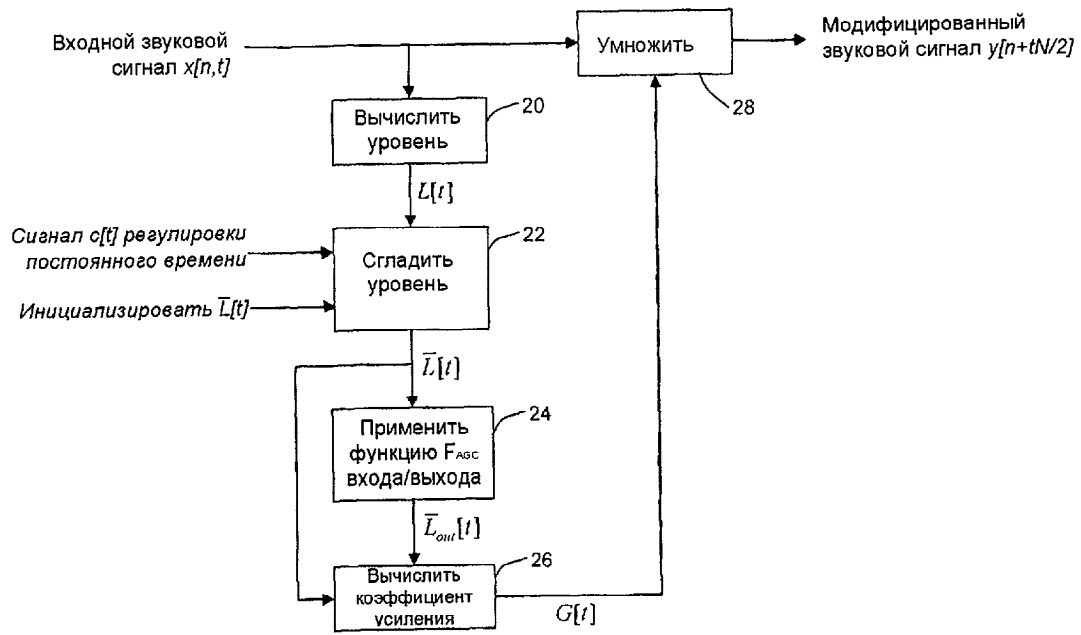
ФИГ.2  
АРУ



ФИГ.4



ФИГ.5



ФИГ.6