

(19)



**Евразийское
патентное
ведомство**

(21) **202193046** (13) **A2**

(12) **ОПИСАНИЕ ИЗОБРЕТЕНИЯ К ЕВРАЗИЙСКОЙ ЗАЯВКЕ**

(43) Дата публикации заявки
2022.06.30

(51) Int. Cl. **G10L 15/00** (2013.01)

(22) Дата подачи заявки
2021.12.03

(54) **АУДИОБЕЙДЖ С ДЕТЕКТОРОМ МЕХАНИЧЕСКИХ КОЛЕБАНИЙ АКУСТИЧЕСКОЙ ЧАСТОТЫ ДЛЯ ВЫДЕЛЕНИЯ РЕЧИ ОПЕРАТОРА**

(31) **2020139758**

(32) **2020.12.03**

(33) **RU**

(71) Заявитель:
**ПУБЛИЧНОЕ АКЦИОНЕРНОЕ
ОБЩЕСТВО "СБЕРБАНК
РОССИИ" (ПАО СБЕРБАНК) (RU)**

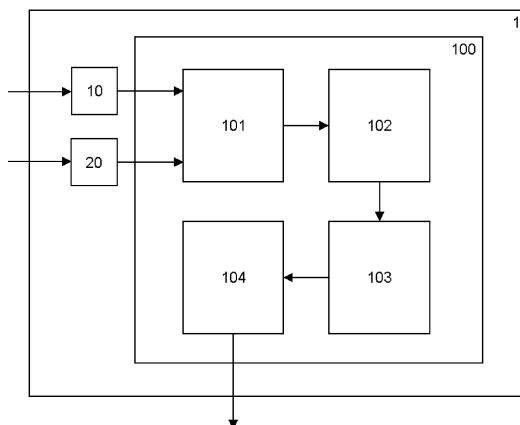
(72) Изобретатель:

**Литвак Юрий Николаевич,
Василенко Алексей Алексеевич,
Цислав Андрей Николаевич, Круглов
Константин Игоревич (RU)**

(74) Представитель:

Герасин Б.В. (RU)

(57) Представленное изобретение относится в общем к измерительной технике, предназначенной для детектирования и записи сигналов, а в частности к устройствам персональной идентификации с функциями аудиозаписи (аудиобейдж) и детектирования механических колебаний акустической частоты для выделения речи оператора на аудиозаписи. Техническим результатом, достигаемым при решении вышеуказанной технической проблемы или технической задачи, является расширение функциональных возможностей аудиобейджа за счет осуществления возможности разметки записываемого цифрового аудиосигнала на основе данных о механических колебаниях акустической частоты. Указанный технический результат достигается благодаря реализации аудиобейджа, содержащего размещенные в едином корпусе и соединенные между собой по меньшей мере один микрофон и вычислительное устройство, при этом вычислительное устройство выполнено с возможностью получения синхронных данных цифрового аудиосигнала и данных механических колебаний; определения характеристического значения для механических колебаний, причем упомянутое характеристическое значение характеризует значение мощности сигнала механических колебаний, возникающих вследствие произнесения звуков речи; разметки данных цифрового аудиосигнала по критерию принадлежности аудиосигнала первому или второму источнику звукового сигнала на основании анализа величины характеристического значения, полученного на предыдущем этапе.



A2

202193046

202193046

A2

АУДИОБЕЙДЖ С ДЕТЕКТОРОМ МЕХАНИЧЕСКИХ КОЛЕБАНИЙ АКУСТИЧЕСКОЙ ЧАСТОТЫ ДЛЯ ВЫДЕЛЕНИЯ РЕЧИ ОПЕРАТОРА.

ОБЛАСТЬ ТЕХНИКИ

[0001] Представленное техническое решение относится, в общем, к измерительной технике, предназначенной для детектирования и записи сигналов, а в частности к устройствам персональной идентификации с функциями аудиозаписи (аудиобейдж) и детектирования механических колебаний акустической частоты для выделения речи оператора на аудиозаписи.

УРОВЕНЬ ТЕХНИКИ

[0002] С научной точки зрения представленное техническое решение относится к устройствам детектирования и записи звука и сопутствующих речи физических признаков, таких как жесты, мимические и иные движения/вибрации тела, связанные с воспроизведением звуков речи, для обнаружения и выделения речи оператора в цифровом сигнале. Данный принцип обнаружения речи применяется в аудио-гарнитурах, в частности Apple AirPods, и раскрыт в заявке US2014093093A1, опубл. 03.04.2014, причем для обнаружения речи в известном решении используется акселерометр обнаружения голосовой активности. Использование датчиков голосовой активности в устройствах, типа гарнитур, позволяет упростить реализацию алгоритмов шумоподавления для звука, воспринимаемого микрофонами, а также улучшить субъективное качество записываемых звуков речи человека, использующего гарнитуру за счет более корректной работы алгоритмов шумоподавления.

[0003] Также известен аудиобейдж, раскрытый в патенте № RU188579 U1, опубл. 17.04.2019. Известный аудиобейдж содержит установленные в корпусе микрофон, блок регистрации, элемент питания и интерфейс, позволяющий передать зарегистрированные сигналы внешнему устройству, два микрофона, разнесенные в корпусе относительно друг друга на заданное расстояние, акселерометр, позволяющий определить ориентацию бейджа, и микроконтроллер, подключенный к акселерометру и блоку регистрации, регистрирующему аудио сигналы по команде микроконтроллера при соответствии данных акселерометра заданной ориентации аудио бейджа.

[0004] Недостатком известных решений является в первом случае – несоответствие аудиогарнитуры форм-фактору средств персональной идентификации (для корпоративного использования), а во втором случае – отсутствие возможности диаризации записываемого сигнала на основании анализа сопутствующих речи оператора механических колебаний поверхности тела.

СУЩНОСТЬ ТЕХНИЧЕСКОГО РЕШЕНИЯ

[0005] Технической проблемой или технической задачей, поставленной в данном техническом решении, является создание нового устройства записи аудиосигнала, обеспечивающего возможность принятия решения о принадлежности фиксируемого в аудиозаписи звукового потока соответствующему диктору (диаризации) и соответствующего разделения записи, например, по каналам стерео, либо в виде дополнительного файла с указанием принадлежности участка записи соответствующему диктору.

[0006] Техническим результатом, достигаемым при решении вышеуказанной технической проблемы или технической задачи, является расширение функциональных возможностей аудиобейджа за счет осуществления возможности разметки записываемого цифрового аудиосигнала на основе данных о механических колебаниях акустической частоты.

[0007] Указанный технический результат достигается благодаря реализации аудиобейджа, содержащего размещенные в едином корпусе и соединенные между собой по меньшей мере один микрофон и вычислительное устройство, при этом вычислительное устройство выполнено с возможностью:

- получения синхронных данных цифрового аудиосигнала и данных механических колебаний;
- определения характеристического значения для механических колебаний, причем упомянутое характеристическое значение характеризует значение мощности сигнала механических колебаний, возникающих вследствие произнесения звуков речи;
- разметки данных цифрового аудиосигнала по критерию принадлежности аудиосигнала первому или второму источнику звукового сигнала на основании анализа величины характеристического значения, полученного на предыдущем этапе.

[0008] В одном из частных примеров реализации айдиобейджа для определения характеристического значения для механических колебаний вычислительное устройство выполнено с возможностью:

- выполнения спектрального анализа данных механических колебаний для получения спектра колебаний, отражающего изменение мощности механических колебаний в зависимости от времени;

- анализа величины мощности, приходящейся на по меньшей мере один участок спектра сигнала механических колебаний, характерный для колебаний, возникающих вследствие произнесения звуков речи;

- выбора величины мощности, приходящейся на по меньшей мере один участок спектра сигнала механических колебаний, в качестве характеристического значения для характеристики механических колебаний.

[0009] В другом частном примере реализации айдиобейджа в качестве спектра сигнала, характерного для колебаний, возникающих вследствие произнесения звуков речи, выбирается диапазон частот от 50 до 300 Гц.

[0010] В другом частном примере реализации айдиобейджа спектральный анализ сигнала механических колебаний выполняется посредством разложения сигнала механических колебаний в ряд Фурье.

[0011] В другом частном примере реализации айдиобейджа по меньшей мере один микрофон реализован на базе МЭМС технологии.

[0012] В другом частном примере реализации айдиобейджа вычислительное устройство выполнено с возможностью разделения данных механических колебаний на временные ряды.

[0013] В другом частном примере реализации айдиобейджа для разметки данных цифрового аудиосигнала по критерию принадлежности аудиосигнала первому или второму источнику звукового сигнала на основании анализа величины характеристического значения сигнала механических колебаний вычислительное устройство выполнено с возможностью:

- сравнения характеристического значения с заранее заданным пороговым значением;

- разметки блока данных цифрового аудиосигнала как относящегося к первому источнику звукового сигнала, если величина характеристического значения сигнала механических колебаний оказывается больше порогового значения;

- разметки блока данных цифрового аудиосигнала как относящемуся к второму источнику звукового сигнала, если величина характеристического значения сигнала механических колебаний оказывается меньше порогового значения.

КРАТКОЕ ОПИСАНИЕ ЧЕРТЕЖЕЙ

[0014] Признаки и преимущества настоящего технического решения станут очевидными из приводимого ниже подробного описания технического решения и прилагаемых чертежей, на которых:

[0015] На Фиг. 1 представлена схема аудиобейджа.

[0016] На Фиг. 2 представлен пример параметров речевого сигнала и механических колебаний.

[0017] На Фиг. 3 подставлен пример общего вида вычислительного устройства.

ОСУЩЕСТВЛЕНИЕ ТЕХНИЧЕСКОГО РЕШЕНИЯ

[0018] Ниже будут описаны понятия и термины, необходимые для понимания данного технического решения.

[0019] В данном техническом решении под системой подразумевается, в том числе компьютерная система, ЭВМ (электронно-вычислительная машина), ЧПУ (числовое программное управление), ПЛК (программируемый логический контроллер), компьютеризированные системы управления и любые другие устройства, способные выполнять заданную, четко определенную последовательность операций (действий, инструкций).

[0020] Под устройством обработки команд подразумевается электронный блок, вычислительное устройство, либо интегральная схема (микропроцессор), исполняющая машинные инструкции (программы).

[0021] Устройство обработки команд считывает и выполняет машинные инструкции (программы) с одного или более устройств хранения данных. В роли устройства хранения данных могут выступать, но не ограничиваясь, жесткие диски (HDD), флеш-память, ПЗУ (постоянное запоминающее устройство), твердотельные накопители (SSD), оптические приводы.

[0022] Программа - последовательность инструкций, предназначенных для исполнения устройством управления вычислительной машины или устройством обработки команд.

[0023] Блок данных - последовательность битов, имеющая фиксированную длину и используемая для представления данных в памяти или для их пересылки.

[0024] Временной ряд — собранный в разные моменты времени статистический материал о значении каких-либо параметров (в простейшем случае одного) исследуемого процесса.

[0025] В соответствии схемой, представленной на Фиг. 1, аудиобейдж 1, в частном случае его реализации, содержит корпус, в котором размещены связанные между собой посредством интерфейсов передачи данных: по меньшей мере один микрофон 1; по меньшей мере один сенсор 20 механических колебаний и вычислительное устройство 100. Упомянутые элементы аудиобейджа 1 образуют единую конструкцию, например, посредством размещения их на единой печатной плате и соединения между собой сборочными операциями, в том числе свинчиванием, пайкой, опрессовкой и пр. В альтернативном варианте реализации заявленного решения сенсор 20 механических колебаний может быть выполнен в виде отдельного выносного элемента, связанного с вычислительным устройством 20 посредством проводной или беспроводной связи.

[0026] В качестве микрофона 1 может быть использован любой известный микрофон, предназначенный для размещения в корпусе портативного устройства, например, выполненный на базе МЭМС технологии. Технология МЭМС (Micro Electro Mechanical Systems - электромеханические микросистемы) — это изготовление электромеханических устройств методами полупроводниковой технологии.

[0027] В качестве сенсора 20 механических колебаний может быть использован, любой сенсор, выполненный с возможностью регистрации механических колебаний акустической частоты, например, акселерометр - прибор, измеряющий проекцию кажущегося ускорения (разность между абсолютным ускорением объекта и гравитационным ускорением, точнее ускорением свободного падения).

[0028] В качестве вычислительного устройства 100 может быть использовано любое известное вычислительное устройство, модифицированное в программно-аппаратной части такими образом, чтобы обеспечить сбор, обработку и хранение данных звукового сигнала. Более подробная конструкция вычислительного устройства 10 будет раскрыта далее в тексте описания со ссылкой на Фиг. 3.

[0029] Для обеспечения возможности разметки записываемого цифрового аудиосигнала на основе данных о механических колебаниях вычислительное устройство 100 может быть оснащено: модулем 101 обработки сигналов, модулем

102 определения характеристик сигнала, модулем 103 разметки аудиосигнала и модулем 104 хранения данных. Перечисленные модули могут быть реализованы на базе программно-аппаратных средств вычислительного устройства 100, выполненные в программной части таким образом, чтобы выполнять приписанные им ниже функции.

[0030] Перед началом работы аудиобейдж 1 размещается вблизи первого источника сигнала, в частности носителя аудиобейджа. Поскольку бейдж предназначен для предоставления информации о его носителе, то он быть может размещен в области груди оператора, который будет являться первым источником сигнала. Соответственно, вторым источником сигнала может быть клиент, находящийся в достаточной близости от аудиобейджа 1 для обеспечения регистрации его голоса. Таким образом, цифровые аудиосигналы, содержащие данные голоса первого или второго источника сигнала, регистрируемые по меньшей мере одним микрофоном 10, поступают в буфер модуля 101 обработки сигналов, в виде потока данных. Из полученных цифровых аудиосигналов модуль 101 формирует массив данных.

[0031] Синхронно с цифровыми аудиосигналами в модуль 101 обработки сигналов также поступают цифровые сигналы с сенсора 20 механических колебаний. Механические колебания, в частности акустической частоты, регистрируемые сенсором 20, могут представлять собой колебания, обусловленные в т.ч. вибрацией поверхности тела в процессе произнесения речи.

[0032] Далее сформированный массив данных цифрового аудиосигнала и данные механических колебаний направляются модулем 101 в модуль 102 определения характеристик сигнала. Примеры данных цифрового аудиосигнала и данных механических колебаний представлены на Фиг. 2, где на верхней схеме изображена функция значения величины звукового давления аудиосигналов первого и второго источника аудиосигнала, зафиксированная одним из микрофонов, а на нижней схеме - механические колебания, регистрируемые в тот же момент времени. Как видно из представленного примера на Фиг. 2 для первого источника аудиосигнала амплитуда механических колебаний будут выше вследствие возникающих вибраций поверхности тела в процессе произнесения речи оператором.

[0033] После получения упомянутым модулем 102 данных цифрового аудиосигнала и сигнала механических колебаний осуществляется определение характеристического значения для механических колебаний. В силу того, что для целей диаризации аудиосигнала требуется выделить только те колебания, природа

происхождения которых связана с произнесением речи, рассчитываемое характеристическое значение должно учитывать периодический характер и характерную частоту (соответствующую частоте основного тона голоса (в процессе произнесения гласных) и частоте модуляции звука (в процессе произнесения звонких согласных)).

[0034] Соответственно, для определения характеристического значения для механических колебаний модуль 102 определения характеристик сигнала выполняет спектральный анализ полученных данных механических колебаний для получения спектра колебаний, отражающего изменение мощности механических колебаний в зависимости от времени. Спектральный анализ сигнала механических колебаний может выполняться известными из уровня техники методами, например, посредством разложения сигнала механических колебаний в ряд Фурье. Далее из полученного спектра колебаний упомянутый модуль 102 извлекает данные спектра сигнала, характерного для колебаний, возникающих вследствие произнесения звуков речи. Например, в качестве спектра сигнала, характерного для колебаний, возникающих вследствие произнесения звуков речи, может быть выбран диапазон частот 50 .. 300 Гц. Соответственно, далее упомянутый модуль 102 выполняет анализ величины мощности, приходящейся на по меньшей мере один участок спектра сигнала механических колебаний, характерного для колебаний, возникающих вследствие произнесения звуков речи, после чего значение мощности сигнала механических колебаний, содержащиеся в данных по меньшей мере одного участка спектра сигнала модулем 102 определения характеристик сигнала выбирается в качестве характеристических значений для механических колебаний.

[0035] В качестве упрощенного варианта реализации процедуры поиска характеристического значения модуль 102 может осуществлять определение величины интеграла модуля сигнала механических колебаний, полученных с частотой дискретизации, соответствующей частоте дискретизации звукового сигнала (8000 или 16000 Гц), на интервале, величиной 10..20 мс после применения к исходной функции ряда оконных фильтров:

$$F_{\text{сигн}_0} = F_{\text{сигн}_0} - \sum_{i=-\Delta/2}^{\Delta/2} \frac{F_{\text{сигн}_i}}{\Delta} - \text{где } \Delta \text{ будет соответствовать интервалу более 100 мс (данное преобразование позволит удалить из сигнала НЧ составляющую);}$$

$F_{\text{сигн}_0} = \sum_{i=-\Delta/2}^{\Delta/2} \frac{F_{\text{сигн}_i}}{\Delta}$ -где Δ будет соответствовать интервалу 3 мс (данное преобразование позволит удалить из сигнала ВЧ составляющую (шумы и сигналы высокой частоты (более 300 Гц))),

где:

$F_{\text{сигн}_0}$ – значение величины сигнала механических колебаний в точке 0 (центральной точке окна фильтра) выбранного/анализируемого временного ряда;

$F_{\text{сигн}_i}$ – значение величины сигнала механических колебаний в точке i выбранного/анализируемого временного ряда (окна фильтра), где i принимает значение от $-\Delta/2$ до $\Delta/2$;

Δ - выбранная ширина окна оконного фильтра в дискретах/сэмплах, например для 20 мс $\Delta = \frac{16000 \text{ сэмпл}}{\text{сек}} * 0,020 \text{ сек} = 320 \text{ сэмпл}$.

[0036] Оставшийся сигнал будет содержать информацию о колебаниях в диапазоне частот в основном 50 .. 300 Гц, что соответствует диапазону частот, соответствующему диапазону в который попадает основной тон голоса. Соответственно интеграл модуля данной функции позволит судить о мощности колебаний в области частот, соответствующей основному тону голоса.

[0037] Дополнительно модуль 102 определения характеристик сигнала может быть выполнен с возможностью разделения данных механических колебаний на несколько временных рядов, причем размер временного ряда может быть задан разработчиком упомянутого модуля 102.

[0038] Далее характеристическое значение для сигнала механических колебаний и данные цифрового аудиосигнала направляются в модуль 103 разметки аудиосигнала, который на основании анализа величины характеристического значения осуществляет разметку данных цифрового аудиосигнала, указывающую на то, к какому источнику аудиосигнала блок данных цифрового аудиосигнала принадлежит. Например, величина характеристического значения сигнала упомянутым модулем 103 могут быть сравнена с заранее заданным пороговым значением и если величина характеристического значения сигнала больше порогового значения, то блок данных цифрового аудиосигнала размечается как относящийся к первому источнику звукового сигнала – носителю бейджа, например, оператору. Если величина характеристического значения сигнала оказывается меньше порогового значения, то блок данных цифрового аудиосигнала размечается как относящийся ко второму источнику звукового сигнала – клиенту 2.

Разметка аудиоданных может осуществляться, как посредством разделения записанного аудиопотока речи дикторов по каналам стерео (речь одного из дикторов – в правый канал, другого – в левый) с последующим сохранением их в виде аудиофайла, либо в виде дополнительного блока данных (отдельного файла либо дополнительной дорожки вышеуказанного аудиофайла) с указанием временных меток, характеризующих время записи реплик одного либо обоих дикторов.

[0039] Полученный аудиофайл с разметкой и/или дополнительный файл разметки аудиоданных может быть сохранен в памяти модуля 104 хранения данных для его передачи в дальнейшем на внешние устройства и системы обработки данных через соответствующие интерфейсы вывода данных, которые будут раскрыты далее в тексте описания.

[0040] В общем виде (см. Фиг. 3) вычислительное устройство (200) содержит объединенные общей шиной информационного обмена один или несколько процессоров (201), средства памяти, такие как ОЗУ (202) и ПЗУ (203) и интерфейсы ввода/вывода (204).

[0041] Процессор (201) (или несколько процессоров, многоядерный процессор и т.п.) может выбираться из ассортимента устройств, широко применяемых в настоящее время, например, таких производителей, как: Intel™, AMD™, Apple™, Samsung Exynos™, MediaTek™, Qualcomm Snapdragon™ и т.п. Под процессором или одним из используемых процессоров в системе (200) также необходимо учитывать графический процессор, например, GPU NVIDIA с программной моделью, совместимой с CUDA, или Graphcore, тип которых также является пригодным для полного или частичного выполнения способа, а также может применяться для обучения и применения моделей машинного обучения в различных информационных системах.

[0042] ОЗУ (202) представляет собой оперативную память и предназначено для хранения исполняемых процессором (201) машиночитаемых инструкций для выполнения необходимых операций по логической обработке данных. ОЗУ (202), как правило, содержит исполняемые инструкции операционной системы и соответствующих программных компонент (приложения, программные модули и т.п.). При этом, в качестве ОЗУ (202) может выступать доступный объем памяти графической карты или графического процессора.

[0043] ПЗУ (203) представляет собой одно или более устройств постоянного хранения данных, например, жесткий диск (HDD), твердотельный накопитель

данных (SSD), флэш-память (EEPROM, NAND и т.п.), оптические носители информации (CD-R/RW, DVD-R/RW, BlueRay Disc, MD) и др.

[0044] Для организации работы компонентов устройства (200) и организации работы внешних подключаемых устройств применяются различные виды интерфейсов В/В (204). Выбор соответствующих интерфейсов зависит от конкретного исполнения вычислительного устройства, которые могут представлять собой, не ограничиваясь: PCI, AGP, PS/2, IrDa, FireWire, LPT, COM, SATA, IDE, Lightning, USB (2.0, 3.0, 3.1, micro, mini, type C), TRS/Audio jack (2.5, 3.5, 6.35), HDMI, DVI, VGA, Display Port, RJ45, RS232 и т.п.

[0045] Конкретный выбор элементов устройства (200) для реализации различных программно-аппаратных архитектурных решений может варьироваться с сохранением обеспечиваемого требуемого функционала.

[0046] Модификации и улучшения вышеописанных вариантов осуществления настоящего технического решения будут ясны специалистам в данной области техники. Предшествующее описание представлено только в качестве примера и не несет никаких ограничений. Таким образом, объем настоящего технического решения ограничен только объемом прилагаемой формулы.

Формула изобретения

1. Аудиобейдж, содержащий размещенные в едином корпусе и соединенные между собой по меньшей мере один микрофон и вычислительное устройство, при этом вычислительное устройство выполнено с возможностью:

- получения синхронных данных цифрового аудиосигнала и данных механических колебаний;
- определения характеристического значения для механических колебаний, причем упомянутое характеристическое значение характеризует значение мощности сигнала механических колебаний, возникающих вследствие произнесения звуков речи;
- разметки данных цифрового аудиосигнала по критерию принадлежности аудиосигнала первому или второму источнику звукового сигнала на основании анализа величины характеристического значения, полученного на предыдущем этапе.

2. Аудиобейдж по п. 1, характеризующийся тем, что для определения характеристического значения для механических колебаний вычислительное устройство выполнено с возможностью:

- выполнения спектрального анализа данных механических колебаний для получения спектра колебаний, отражающего изменение мощности механических колебаний в зависимости от времени;
- анализа величины мощности, приходящейся на по меньшей мере один участок спектра сигнала механических колебаний, характерного для колебаний, возникающих вследствие произнесения звуков речи;
- выбора величины мощности, приходящейся на по меньшей мере один участок спектра сигнала механических колебаний, в качестве характеристического значения для характеристики механических колебаний.

3. Аудиобейдж по п. 2, характеризующийся тем, что в качестве спектра сигнала, характерного для колебаний, возникающих вследствие произнесения звуков речи, выбирается диапазон частот от 50 до 300 Гц.

4. Аудиобейдж по п. 2, характеризующийся тем, что спектральный анализ сигнала механических колебаний выполняется посредством разложения сигнала механических колебаний в ряд Фурье.

5. Аудиобейдж по п. 1, характеризующийся тем, что по меньшей мере один микрофон реализован на базе МЭМС технологии.

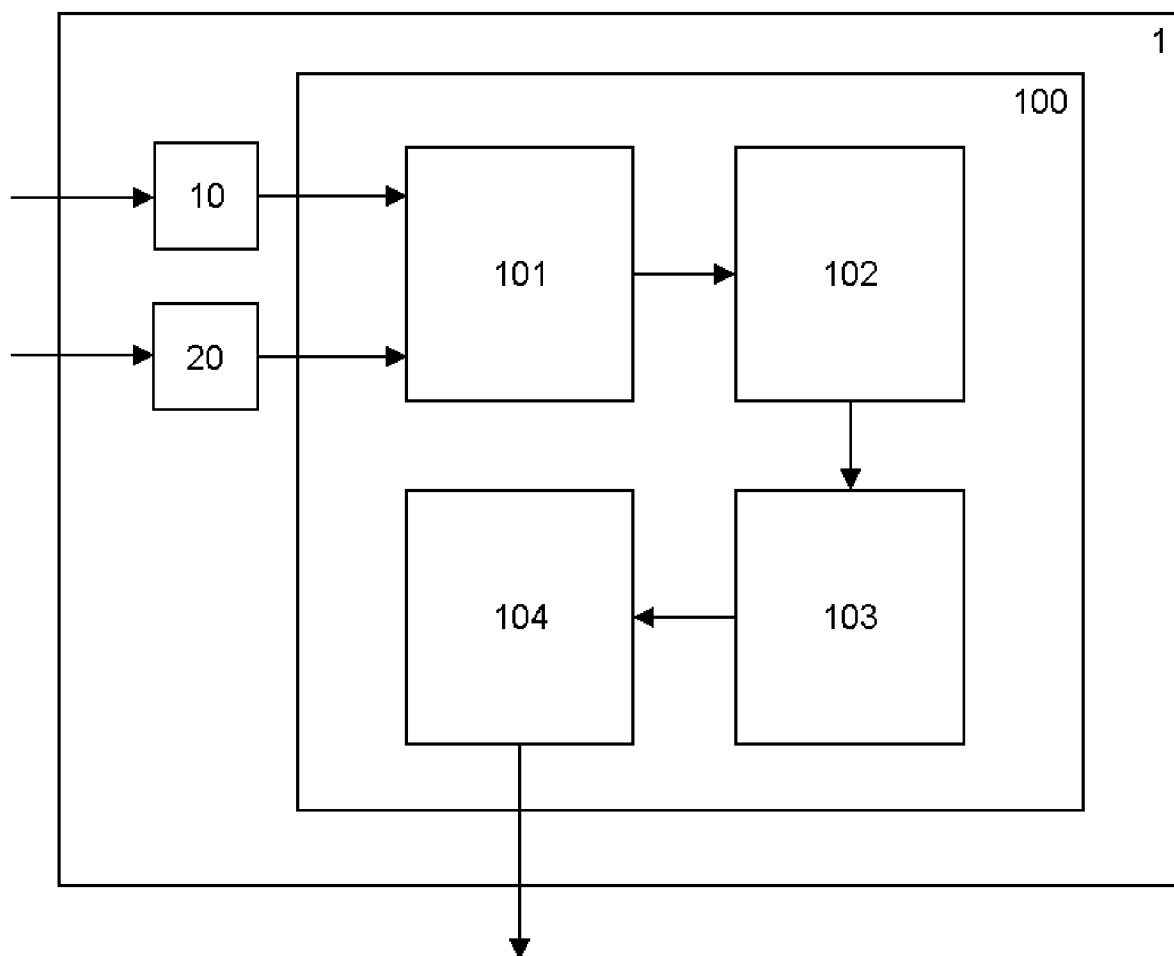
6. Аудиобейдж по п. 1, характеризующийся тем, что вычислительное устройство выполнено с возможностью разделения данных механических колебаний на временные ряды.

7. Аудиобейдж по п. 1, характеризующийся тем, что для разметки данных цифрового аудиосигнала по критерию принадлежности аудиосигнала первому или второму источнику звукового сигнала на основании анализа величины характеристического значения сигнала механических колебаний вычислительное устройство выполнено с возможностью:

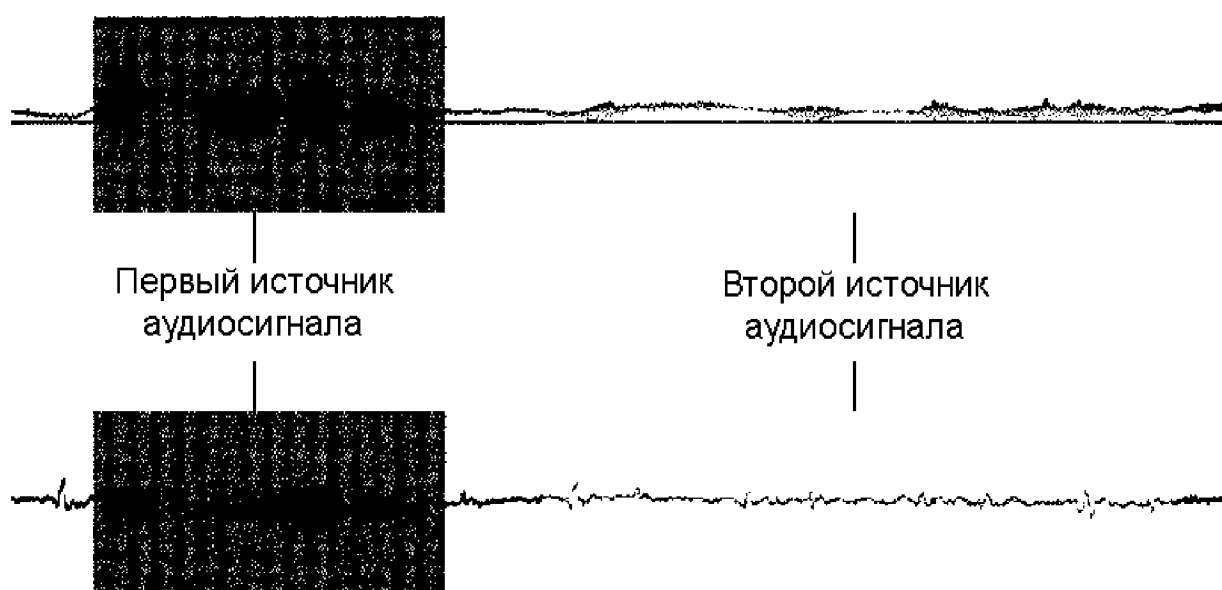
- сравнения характеристического значения с заранее заданным пороговым значением;

- разметки блока данных цифрового аудиосигнала как относящегося к первому источнику звукового сигнала, если величина характеристического значения сигнала механических колебаний оказывается больше порогового значения;

- разметки блока данных цифрового аудиосигнала как относящегося к второму источнику звукового сигнала, если величина характеристического значения сигнала механических колебаний оказывается меньше порогового значения.



ФИГ. 1



ФИГ. 2

ФИГ. 3

