



(51) МПК
G10L 19/02 (2013.01)
G10L 15/02 (2006.01)
G10L 21/14 (2013.01)

ФЕДЕРАЛЬНАЯ СЛУЖБА
 ПО ИНТЕЛЛЕКТУАЛЬНОЙ СОБСТВЕННОСТИ

(12) ОПИСАНИЕ ИЗОБРЕТЕНИЯ К ПАТЕНТУ

(21)(22) Заявка: 2014134745/08, 26.08.2014

(24) Дата начала отсчета срока действия патента:
 26.08.2014

Приоритет(ы):

(22) Дата подачи заявки: 26.08.2014

(43) Дата публикации заявки: 20.03.2016 Бюл. № 8

(45) Опубликовано: 10.07.2016 Бюл. № 19

(56) Список документов, цитированных в отчете о поиске: RU 2 234 746 C2, 20.08.2004. RU 2 199 157 C2, 20.02.2003. RU 2 103 753 C1, 27.01.1998. RU 2 376 656 C1, 20.12.2009. US 2013/0325154 A1, 05.12.2013.

Адрес для переписки:

143026, Москва, территория инновационного центра "Сколково", ул. Луговая, 4, ООО "Центр интеллектуальной собственности "Сколково"

(72) Автор(ы):

Кнеллер Эммануил Григорьевич (RU),
 Караульных Денис Владимирович (RU)

(73) Патентообладатель(и):

Общество с Ограниченной ответственностью
 "ИстраСофт" (RU)

(54) СИСТЕМА И СПОСОБ ПЕРЕВОДА РЕЧЕВОГО СИГНАЛА В ТРАНСКРИПЦИОННОЕ ПРЕДСТАВЛЕНИЕ С МЕТАДАННЫМИ

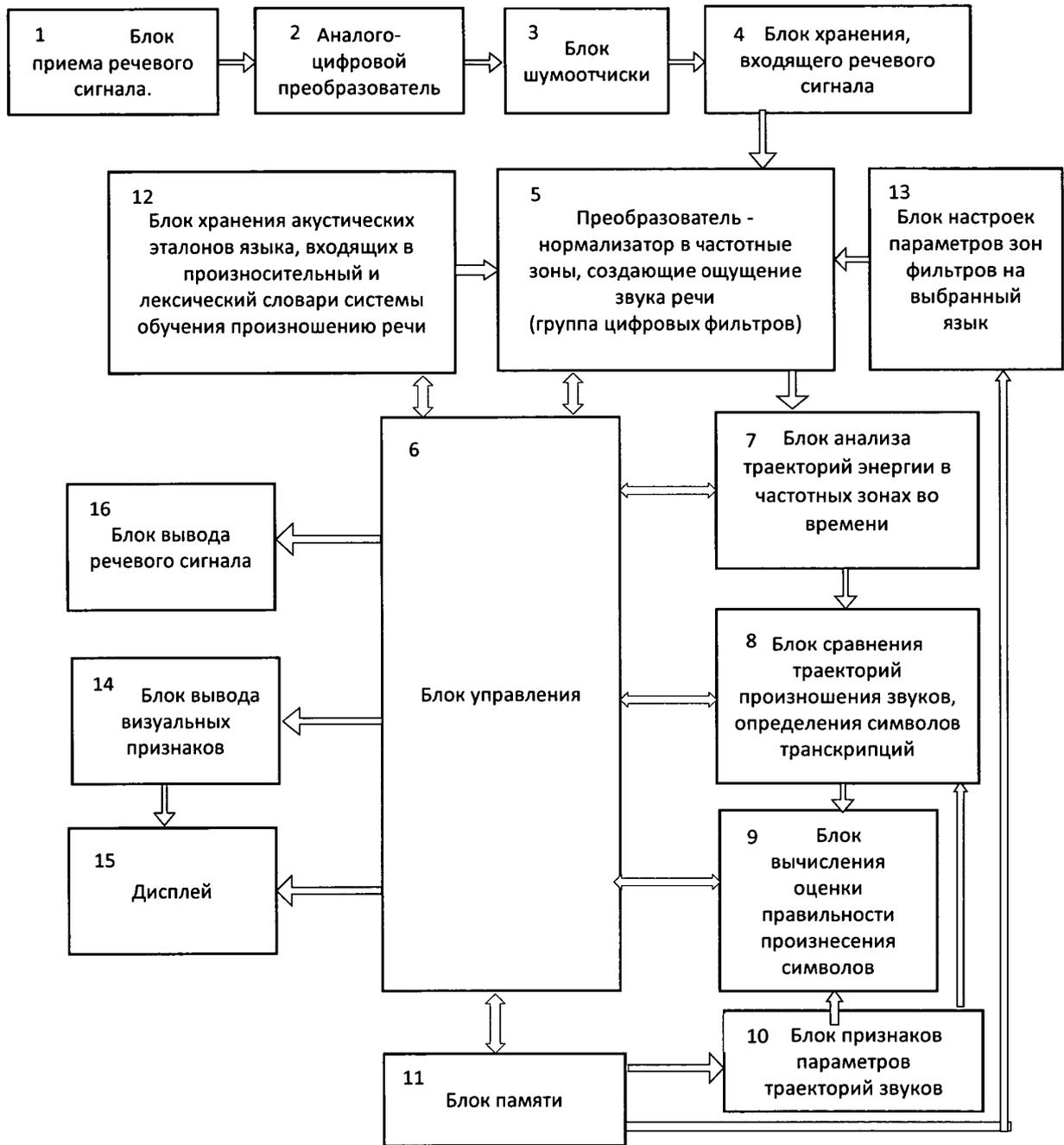
(57) Реферат:

Данное изобретение относится к технологиям распознавания речи, т.е. перевода звукового сигнала, содержащего речь, в транскрипционное представление. Технический результат заключается в повышении точности распознавания звуков речи. Производят прием и предварительную обработку речевого сигнала, далее преобразуют речевой сигнал группой цифровых фильтров в частотные зоны, превышение относительной энергии в которых

во времени создает ощущение звуков языка, затем нормализуют полученные частотные зоны с учетом характеристик чувствительности улитки уха по частоте, амплитуде и времени, после чего определяют акустические признаки речевого сигнала и типы звуков, в конечном итоге нормализуют временные параметры и отображают обработанный речевой сигнал. 2 н. и 16 з.п. ф-лы, 7 ил.

RU 2 589 851 C 2

RU 2 589 851 C 2



Фиг. 1



FEDERAL SERVICE
FOR INTELLECTUAL PROPERTY

(51) Int. Cl.
G10L 19/02 (2013.01)
G10L 15/02 (2006.01)
G10L 21/14 (2013.01)

(12) **ABSTRACT OF INVENTION**

(21)(22) Application: 2014134745/08, 26.08.2014

(24) Effective date for property rights:
26.08.2014

Priority:

(22) Date of filing: 26.08.2014

(43) Application published: 20.03.2016 Bull. № 8

(45) Date of publication: 10.07.2016 Bull. № 19

Mail address:

143026, Moskva, territorija innovatsionnogo tsentra
"Skolkovo", ul. Lugovaja, 4, OOO "TSentr
intellektualnoj sobstvennosti "Skolkovo"

(72) Inventor(s):

**Kneller Emmanuil Grigorevich (RU),
Karaulnykh Denis Vladimirovich (RU)**

(73) Proprietor(s):

**Obshchestvo s Ogranichennoj otvetstvennostyu
"IstraSoft" (RU)**

(54) **SYSTEM AND METHOD OF CONVERTING VOICE SIGNAL INTO TRANSCRIPT PRESENTATION WITH METADATA**

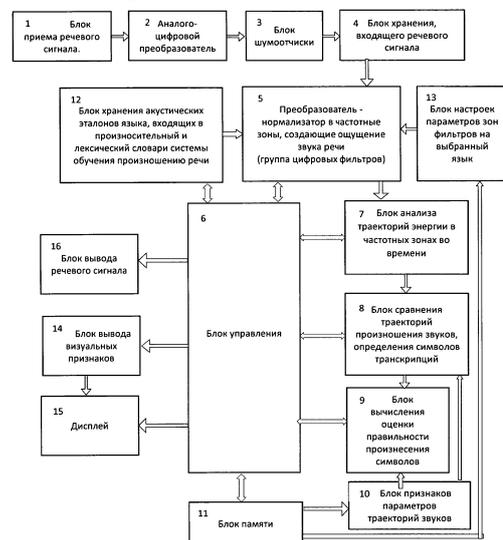
(57) Abstract:

FIELD: information technology.

SUBSTANCE: present invention relates to techniques for speech recognition, that is, conversion of audio signal containing speech into a transcript representation. Method comprises receiving and preliminary processing of speech signal, converting speech signal group of digital filters into frequency areas, exceeding of relative energy in time of which creates a perception of language sounds, then standardising obtained frequency area taking into account characteristics of sensitivity of volute ear frequency, amplitude and time, after which acoustic criteria of speech signal and types of sounds, finally standardising time parameters and displaying a processed voice signal.

EFFECT: technical result consists in improvement of accuracy of recognition of speech sounds.

18 cl, 7 dwg



Фиг. 1

RU 2 589 851 C2

RU 2 589 851 C2

ОБЛАСТЬ ТЕХНИКИ

Данное изобретение относится к технологиям распознавания речи, т.е. перевода звукового сигнала, содержащего речь, в транскрипционное представление, состоящую из символов отдельных звуков, транскрипции слов, предложений и их метаданных, визуализацию физических характеристик речевого сигнала, создающих ощущение звуков речи, оценку правильности произношения, входящих в произносительный словарь системы обучения произношению речи.

УРОВЕНЬ ТЕХНИКИ

Известен способ, описанный в заявке на патент RU 2234746 «СПОСОБ ДИКТОРОНЕЗАВИСИМОГО РАСПОЗНАВАНИЯ ЗВУКОВ РЕЧИ», опубликовано 20.08.2004, патентообладатель «Пермский государственный университет». Его использование позволяет получить технический результат в виде повышения вероятности правильного распознавания звуков речи. Способ включает в себя сегментацию речевого сигнала по времени, определение периодичности каждого звукового сегмента для соотнесения звукового сегмента к конкретному виду звуков речи, определение амплитуды и частоты каждой из первых трех формант в спектре звукового сегмента в качестве информативных признаков, интеграция этих признаков для каждого звукового сегмента, фонемное распознавание каждого звукового сегмента путем сопоставления его интегральных значений, принятие решения о распознаваемом звуке речи и представление его в виде буквенного обозначения. Технический результат достигается благодаря тому, что соотнесение звукового сегмента осуществляют к голосовому, шумному или шумно-голосовому виду звуков речи, выполняют основную сегментацию речевого сигнала по трем основным режимам, при фонемном распознавании сопоставляют интегральные значения информативных признаков каждого звукового сегмента в зависимости от числа формант в звуковом сегменте, устанавливают временные границы звуков речи в зависимости от изменения фонемной принадлежности звукового сегмента, после чего и принимают решение относительно распознаваемого звука речи.

Недостатком данного технического решения является низкая различительная способность, связанная с тем, что форманты не являются теми характеристиками, которые определяют звук, - это только один из способов создания в речевом сигнале соответствующих физических характеристик, создающих ощущение звука. Известно, например, что звук А может иметь от одной до трех формант в зависимости от голоса говорящего.

Известен способ, описанный в заявке на патент RU2297676 «СПОСОБ РАСПОЗНАВАНИЯ СЛОВ В СЛИТНОЙ РЕЧИ», опубликовано 20.04.2007, патентообладатель Федеральное государственное научное учреждение научно-исследовательский институт "Специализированные вычислительные устройства защиты и автоматика"

Изобретение относится к автоматике и вычислительной технике. Его использование в системах управления технологическим, бытовым и другим оборудованием, в автоматических справочных системах, системах автоматического перевода, системах понимания речи и т.д. позволяет получить технический результат в виде повышения точности распознавания слов в слитной речи. Способ состоит в том, что с произнесением речевого высказывания периодически берут выборки акустического сигнала этого высказывания, оцифрованного с заданной частотой квантования, через фиксированные интервалы времени и по совокупности этих выборок вычисляют функционал, определяющий текущее акустическое состояние, при этом полученную

последовательность текущих акустических состояний используют для восстановления последовательности слов (рабочей гипотезы), произнесенных в исходном речевом высказывании, для чего применяют сеть лексического декодирования, которая задает закономерности следования эталонных акустических состояний в языке. Технический результат достигается за счет того, что проводится поиск рабочей гипотезы, являющийся оптимальным в смысле максимума степени ее совпадения с исходным речевым сигналом, что обеспечивается использованием алгоритма перемещаемого маркера, при этом рабочую гипотезу восстанавливают из маркера, который в этот момент времени находится в конечной вершине сети лексического декодирования.

Несмотря на то, что в данном способе различительная способность выше, чем в предыдущем способе, однако, аналогично с предыдущим известным способом распознавания речи к числу недостатков данного метода следует также отнести длительность процесса распознавания, обусловленного необходимостью обращения к эталонным образцам, а также поочередным распознаванием каждой форманты в слове.

СУЩНОСТЬ ИЗОБРЕТЕНИЯ

Данное изобретение направлено на устранение недостатков, присущих существующим решениям.

Техническим результатом данного изобретения является распознавание звуков речи с высокой точностью, в реальном времени, вне зависимости от голоса говорящего, его индивидуальных особенностей, шума окружающей среды, специфических особенностей разных языковых систем, основанный на выделении и измерении только тех физических характеристик акустического сигнала, которые непосредственно создают ощущения звуков данного языка.

Способ перевода речевого сигнала в транскрипционное представление с метаданными реализуется согласно изобретению следующим образом:

Производят прием и предварительную обработку речевого сигнала, далее преобразуют речевой сигнал группой цифровых фильтров в частотные зоны, превышение относительной энергии в которых во времени создает ощущение звуков языка, затем нормализуют полученные частотные зоны с учетом характеристик чувствительности улитки уха по частоте, амплитуде и времени, после чего определяют акустические признаки речевого сигнала и типы звуков, в конечном итоге нормализуют временные параметры и отображают обработанный речевой сигнал.

В некоторых вариантах осуществления изобретения, принятый речевой сигнал очищают от шумов и обрабатывают аналого-цифровым преобразователем с предустановленной частотой дискретизации.

В некоторых вариантах осуществления изобретения, на этапе распознавания в исходном речевом сигнале одновременно с получением параметров траекторий из акустических эталонов формируются последовательности символов транскрипции, звука, слова, предложения, визуализация признаков звуков и оценка их правильности произношения относительно акустических эталонов.

В некоторых вариантах осуществления изобретения определение комбинаторного набора акустических признаков сегмента речевого сигнала осуществляют параллельно и одномоментно.

В альтернативных вариантах осуществления изобретения визуализацию значимых параметров выводят в виде понятных 2D и 3D графических представлений, показывающих отличие оцениваемых характеристик эталонного произношения от произносимого.

В некоторых вариантах осуществления изобретения, что при определении

акустических признаков речевого сигнала анализируют данные частотные зоны и выделяют группы звуков по похожим траекториям энергий.

В некоторых вариантах осуществления изобретения осуществляют формирование словаря на основе размеченных траекторий энергии зон, определяющих звуки, слова, предложения, полученные из акустических эталонов языка, входящих в произносительный словарь системы обучения произношению речи.

В некоторых вариантах осуществления изобретения звук определяется превышением относительной энергией в одной, двух или нескольких значащих зонах на временном отрезке сигнала в пределах 30-50 мс.

В некоторых вариантах осуществления изобретения звуки по временным параметрам определяются как статические, если количество энергетических зон, создающих ощущения звуков, не изменяется в течение 30-100 мс и более.

В некоторых вариантах осуществления изобретения звуки по временным параметрам определяются как взрывные, если предшествует отсутствие превышение энергии во всех значащих зонах на длине 5-25 мс, затем скачок энергии в большой группе фильтров, занимающий 15-20 мс, с последующим переходом в зоны, определяющие звук.

В некоторых вариантах осуществления изобретения для определения амплитудно-модулированных звуков измеряют частоту модуляции и отклики фильтров.

В некоторых вариантах осуществления изобретения для определения интонации измеряют относительное изменение траектории частоты основного тона, на основе которой определяют тип интонации, на длине звуков, слов или предложений.

В некоторых вариантах осуществления изобретения для определения акцента и ритмики сравнивают относительное время звучания звуков в слове относительно эталонного.

В некоторых вариантах осуществления изобретения для визуализации оцениваемых характеристик эталонного произношения от произносимого осуществляют нормализацию по времени произнесения выводимого сигнала относительно эталонного.

В некоторых вариантах реализации изобретение представляет собой распределенную систему перевода речевого сигнала, в транскрипционное представление с метаданными, включающую:

- i. одно или более устройств обработки команд;
- ii. одно или более устройств хранения данных;
- iii. одну или более программ, реализующих вышеуказанный способ перевода речевого сигнала в транскрипционное представление с метаданными, причем шаги данного способа могут выполняться на разных компьютерных системах.

КРАКТОЕ ОПИСАНИЕ ИЗОБРАЖЕНИЙ

Признаки и преимущества настоящего изобретения станут очевидными из приводимого ниже подробного описания изобретения и прилагаемых чертежей, на которых:

Фиг. 1 - структурная схема системы транскрибации речи.

Фиг. 2 - функциональная схема выделения признаков транскрипции, включая этапы преобразования в частотные зоны, анализ траекторий, энергий в них, сравнение с эталонными, оценка соответствия объединенных признаков звукам и формирования последовательности символов, обозначающих группы фонем и ее декодирования в текст.

Фиг. 3 - параметры спектральных зон речевого сигнала, создающие ощущения звуков речи.

Фиг. 4 - графическое представление характеристик речевого сигнала для визуального

сравнения правильности произношения относительно эталона, оценка производится на сравнении параметров интегрированной энергии за время звучания звука в зонах, соответствующих звуков.

Фиг. 5 - схема математической модели артикуляционного аппарата.

5 ПОДРОБНОЕ ОПИСАНИЕ ИЗОБРЕТЕНИЯ

Данное изобретение в различных своих вариантах осуществления может быть выполнено в виде способа, реализуемого на компьютере, в виде системы или машиночитаемого носителя, содержащего инструкции для выполнения вышеупомянутого способа.

10 В некоторых вариантах реализации изобретение может быть реализовано в виде распределенной компьютерной системы.

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

Под устройством обработки команд подразумевается электронный блок либо интегральная схема (микропроцессор), исполняющая машинные инструкции (программы).

20 Устройство обработки команд считывает и выполняет машинные инструкции (программы) с одного или более устройства 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.

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

Ниже будут описаны понятия и определения, необходимые для подробного раскрытия осуществляемого изобретения.

30 Метаданные - информация об интонации, акценте, ритмике, тоне (мелодики речи), громкости, темпе речи и ее отдельных отрезков, особенностей фонации.

Звуки классифицируются по временным параметрам.

Статические звуки - энергетика зон незначительно изменяются в течение 30-100 мс (миллисекунда).

35 Динамические звуки - энергетика зон изменяются по определенным траекториям в течение 30-100 мс (миллисекунда).

Взрывные звуки - отсутствие энергии на длине 5-25 мс (миллисекунда), затем скачок энергии в группе фильтров, занимающий 15-20 мс (миллисекунда), с последующим переходом в зоны определяющие звук.

40 Произносительный словарь - это словарь фонетических разборов, в котором каждое слово рассматривается с точки зрения его фонетической структуры, учитывая метаданные.

Спектр сигнала - в радиотехнике это результат разложения сигнала на более простые частотные зоны в базисе ортогональных функций.

45 Согласно изобретению способ перевода речевого сигнала в транскрипционное представление с метаданными, включает следующие шаги:

производят прием и предварительную обработку речевого сигнала:

Речевой сигнал в виде звукового потока данных поступает на вход блока приема речевого сигнала (Фиг. 1) системы распознавания речи и передается далее на вход

аналого-цифрового преобразователя, где осуществляют преобразование сигнала в цифровой вид. Полученный цифровой речевой сигнал очищают от шумов и сохраняют для дальнейшей обработки, передают на вход преобразователя - нормализатора;

преобразуют речевой сигнал группой цифровых фильтров в частотные зоны, превышение относительной энергии в которых во времени создает ощущение звуков языка:

С помощью преобразователя - нормализатора осуществляют спектральное преобразование группой цифровых фильтров с параметрами преобразования, нормализованными с учетом характеристик чувствительности улитки уха по частоте, амплитуде и времени, выбор отсечки частотных полос фильтров определяют на основании частотных зон, превышение относительной энергии в которых создает ощущения звуков соответствующего языка;

определяют акустические признаки речевого сигнала и типы звуков:

Акустические признаки речевого сигнала, характерные для групп фонем, используемых в качестве базовых элементов для распознавания, основаны на физических принципах генерации звуков артикуляционным аппаратом человека, поясняются на схеме математической модели артикуляционного аппарата (Фиг. 5). Возбуждающий сигнал, тон, шум или их комбинация, проходя через модулятор, резонаторы и ключи, управляемые мозгом, изменяют распределение энергии в частотных зонах в зависимости от требуемого звука, при этом положение резонаторов артикуляционных органов для многих пар звонких - глухих звуков типа [А]-[Х], [З]-[С], одинаковое, что вызывает одинаковые отклики в зонах фильтров, отличие заключается только в типе возбуждающей функции, что позволяет при одинаковых откликах наряду с другими признаками классифицировать звуки, превышение относительной энергии в одной или двух значащих частотных зонах, ширина и положение которых в частотном спектре на значимом отрезке времени определяет звук языка (Фиг. 3, 3а, 3б).

Акустический сигнал ощущается как звук речи, если длительность характеристик, определяющих звук не менее 30-100 мс, при меньших или больших значениях теряется ощущение звука, то есть мы слышим акустический сигнал, но не можем определить, что это за звук. Экспериментально установлено, что при произношении слова, предложения, звуки в слове имеют постоянные временные пропорции к общему времени его произнесения, то есть время произнесения каждого звука имеет постоянную относительную длину, говорим быстро или медленно соотношение звуков в процентном отношении к общему времени остается постоянным для одного типа произношения данного языка. Время произнесения звука является одним из основных параметров, влияющих на ощущение звуков. Обычно, когда произносится не тоновый звук в непрерывной речи, например не тоновый звук [t], остановка и запуск связок занимает намного дольше времени и энергии. Связки продолжают генерировать. Обозначают звук [t] только временем, хотя реально в это время связки генерируют звук, а воспринимается он как не тоновый звук [t]. Изменением во времени звука создаются и другие ощущения речи, так ударный слог имеет относительное время произнесения в два раза большее, чем безударный. Уменьшение времени паузы перед взрывным звуком меняет ощущение звука, например звук [к] мы ощутим как звук [г]. Изменение времени звучания звуков в слове от принятого в данном регионе воспринимается как региональный говор или акцент.

Как указано выше, для определения акустических признаков речевого сигнала используют как исходный речевой сигнал, так и эталонный. На основе комбинаций значений акустических признаков определяется группа фонем, к которой относится

речевой сигнал, после чего производится оценка их соответствия эталонному и визуализируются признаки, позволяющие увидеть отличие физических характеристик, на основании которых даются рекомендации обучаемому в рамках текущего шага обучения (Фиг. 4).

5 При этом процедура выделения акустических признаков, применяемых для определения групп фонем, из сигнала согласно изобретению осуществляется следующим образом: сканируют выходы фильтров, находят глобальный и локальные максимумы энергии, затем в пределах ± 30 мс (миллисекунда) от максимума анализируют изменения энергии в зонах, определяют временной тип звука, то есть статический, динамический
10 или взрывной, находят близкие по траекториям отклики в эталоне, по которым идентифицируют тип фонемы, сканируют дальше и ищут следующий локальный максимум и процедуру повторяют дальше;

При классификации групп фонем в рамках каждого окна локального энергетического максимума речевого сигнала параллельно и одновременно программно-аппаратным
15 образом используют следующий набор акустических признаков, позволяющих однозначно определить каждую из рассматриваемых в рамках изобретения групп признаков фонем: превышение относительной энергии в частотных зонах выделенных цифровыми фильтрами зонах, в том числе отдельной зоны определяющей наличие/
отсутствие основного тона, время воздействия, изменения динамических характеристик
20 энергии и пауз в зонах, наличие/отсутствие пауз и признаков взрывного звука, анализ речевого сигнала в каждом окне осуществляют на наличие/отсутствие акустических признаков, присущих группам фонем в речевом сигнале;

нормализуют временные параметры и отображают обработанный речевой сигнал;

В процессе обучения из произносительного словаря системы обучения произношению
25 речи, в соответствии с методикой обучения выбирается и произносится, звук, слово или фраза. Обучаемый старается произнести услышанный им звуковой фрагмент с таким же темпом, интонацией и акцентом. Произнесенный фрагмент нормализуется по времени. Все характеристики этого фрагмента как эталонного, так и произносимого обучаемым,
выводятся в виде звукового и графического представления (Фиг. 4), позволяющего
30 визуально сравнить правильность произношения характеристик звукового объекта относительно эталона, получить оценку проклассифицированных на основе данного изобретения признаков соответствующих звуков. При несоответствии произнесения звукового фрагмента эталону определяется отклонение характеристики, и при их
наличии даются рекомендации в соответствии с методикой обучения по их устранению.

35 Специалисту в данной области очевидно, что конкретные варианты осуществления способа и системы перевода речевого сигнала в транскрипционное представление с метаданными были описаны здесь в целях иллюстрации, допустимы различные модификации, не выходящие за рамки и сущности объема изобретения.

40 Формула изобретения

1. Способ перевода речевого сигнала в транскрипционное представление с метаданными, включающий следующие шаги:

- производят прием и предварительную обработку речевого сигнала;
- преобразуют речевой сигнал группой цифровых фильтров в частотные зоны;
- 45 - нормализуют полученные частотные зоны с учетом характеристик чувствительности улитки уха по частоте, амплитуде и времени;
- определяют акустические признаки речевого сигнала и типы звуков;
- нормализуют временные параметры принятого речевого сигнала и отображают

его после обработки.

2. Способ по п. 1, характеризующийся тем, что принятый речевой сигнал очищают от шумов и обрабатывают аналого-цифровым преобразователем с фиксированной частотой дискретизации.

5 3. Способ по п. 1, характеризующийся тем, что цифровой фильтр может быть программным и/или аппаратным.

4. Способ по п. 1, характеризующийся тем, что при преобразовании речевого сигнала в частотные зоны формируют последовательности символов транскрипции звука, слова, предложения и производят визуализацию параметров звуков и оценку их правильности
10 произношения относительно акустических эталонов.

5. Способ по п. 1, характеризующийся тем, что отображают акустические параметры звука в виде 2D и 3D графических представлений, позволяющие сравнить правильность произношения характеристик звукового объекта относительно акустического эталона.

6. Способ по п. 1, характеризующийся тем, что отображают превышение
15 относительной энергии по времени речевого сигнала в режиме реального времени.

7. Способ по п. 1, характеризующийся тем, что звуки речи классифицируются в зависимости от физических параметров генерации на тоновые, шумовые, амплитудно-модулированные, частотно-модулированные, взрывные и паузы.

8. Способ по п. 1, характеризующийся тем, что при определении типа звука зона
20 тона не используется.

9. Способ по п. 1, характеризующийся тем, что при преобразовании речевого сигнала группой цифровых фильтров звук определяется превышением относительной энергии по крайней мере в одной зоне на временном отрезке сигнала в пределах 30-50 мс (миллисекунда).

25 10. Способ по п. 1, характеризующийся тем, что при определении акустических признаков речевого сигнала сравнивают принятый речевой сигнал с эталонным.

11. Способ по п. 1, характеризующийся тем, что при определении акустических признаков речевого сигнала звуки по временным параметрам определяются как статические, если количество энергетических зон, создающих ощущения звуков, не
30 изменяется в течение 30-100 мс (миллисекунда) и более.

12. Способ по п. 1, характеризующийся тем, что при определении акустических признаков речевого сигнала звуки по временным параметрам определяются как динамические, если количество энергетических зон, создающих ощущения звуков, изменяется по определенным траекториям в течение 30-100 мс (миллисекунда).

35 13. Способ по п. 1, характеризующийся тем, что при определении акустических признаков речевого сигнала звуки по временным параметрам определяются как взрывные, если энергия отсутствует на длине 5-25 мс (миллисекунда), затем происходит скачок энергии в группе фильтров, занимающий 15-20 мс (миллисекунда), с последующим переходом в зоны, определяющие звук.

40 14. Способ по п. 1, характеризующийся тем, что при определении акустических признаков речевого сигнала для определения амплитудно-модулированных звуков типа [з], [ж] измеряют частоту модуляции и отклики фильтров.

15. Способ по п. 1, характеризующийся тем, что при определении акустических признаков речевого сигнала для определения интонации измеряют относительное
45 изменение траектории частоты основного тона, на основе которой определяют тип интонации, на длине звуков, слов или предложений.

16. Способ по п. 1, характеризующийся тем, что при определении акустических признаков речевого сигнала для определения акцента и ритмики сравнивают

относительное время звучания звуков в слове относительно эталонного.

17. Способ по п. 1, характеризующийся тем, что при определении акустических признаков речевого сигнала на основании отличий физических характеристик принятого речевого сигнала от эталонных дают рекомендации обучаемому в рамках текущего шага обучения.

18. Система перевода речевого сигнала в транскрипционное представление с метаданными включает:

- одно или более устройств хранения информации;
- средства, обеспечивающие сетевое взаимодействие с удаленными компьютерными системами;
- по крайней мере один дисплей;
- по крайней мере одно устройство ввода;
- один или более процессоров, где одна или более компьютерных программ считываются с одного и более устройств хранения информации и выполняются на одном и более процессорах, причем одна или более программ содержат инструкции для выполнения способа по п. 1.

20

25

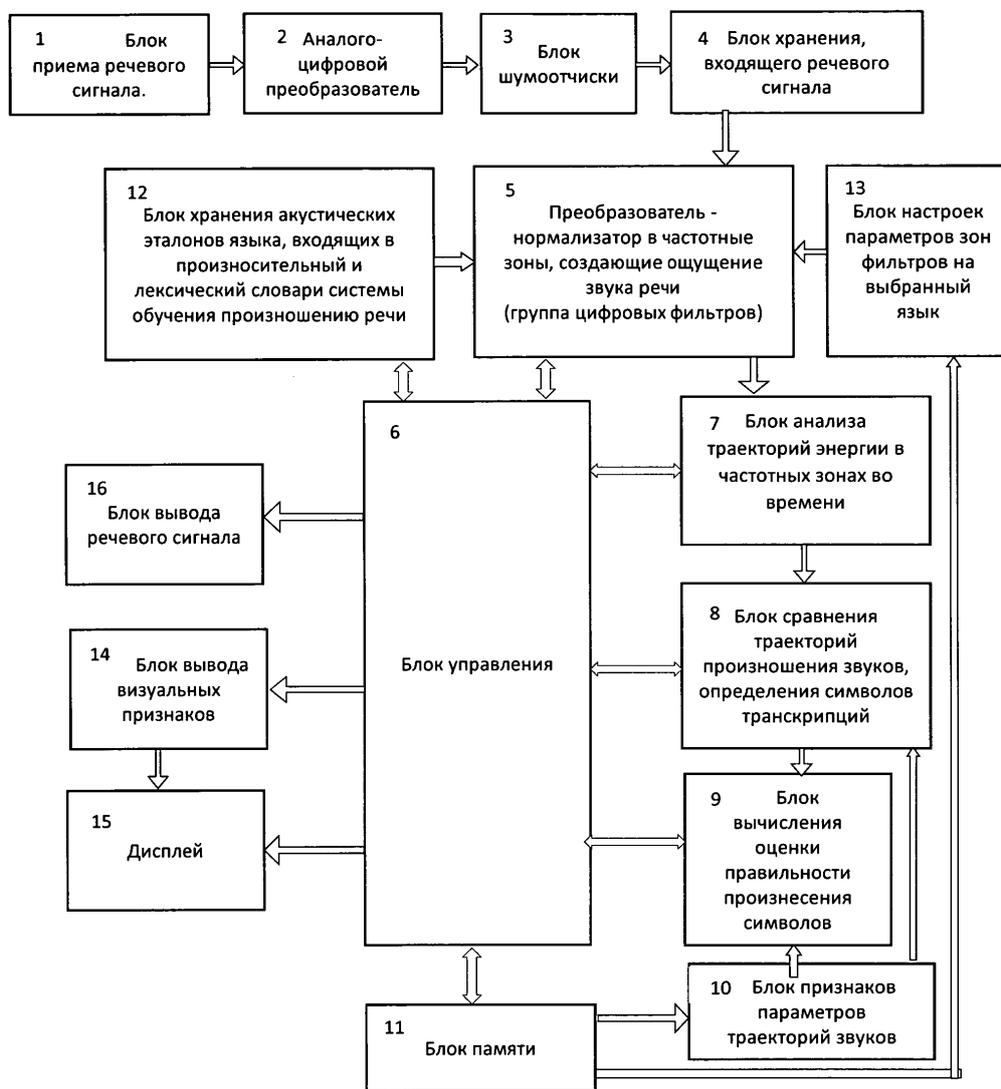
30

35

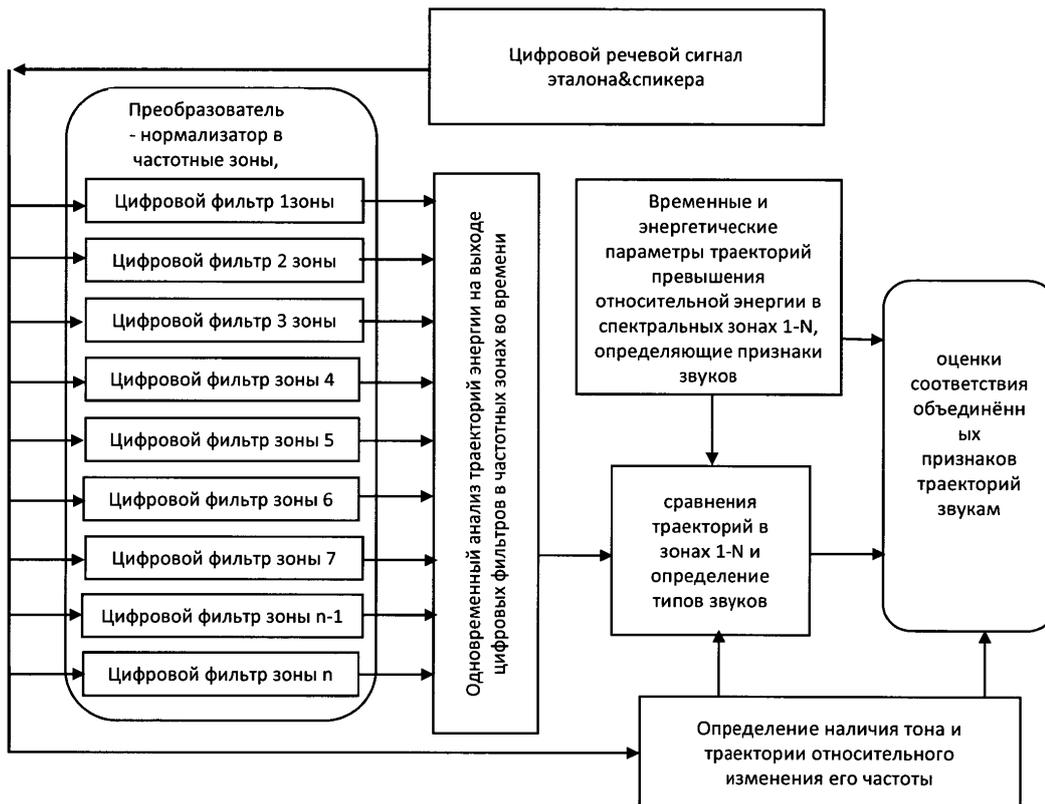
40

45

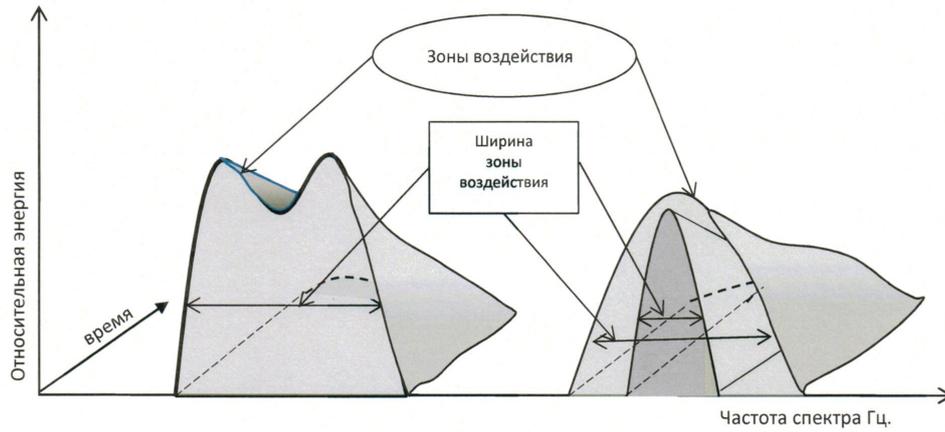
Структурная схема системы транскрибации речи



Фиг. 1



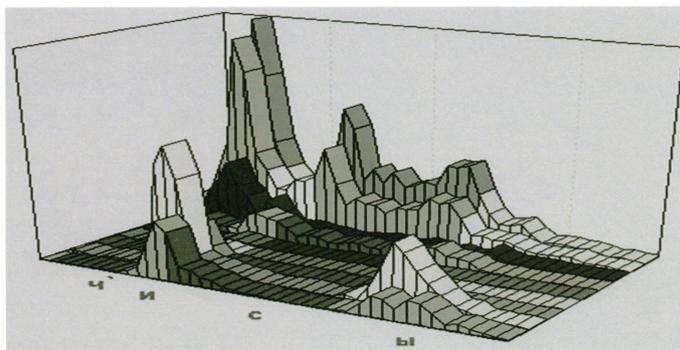
Фиг. 2



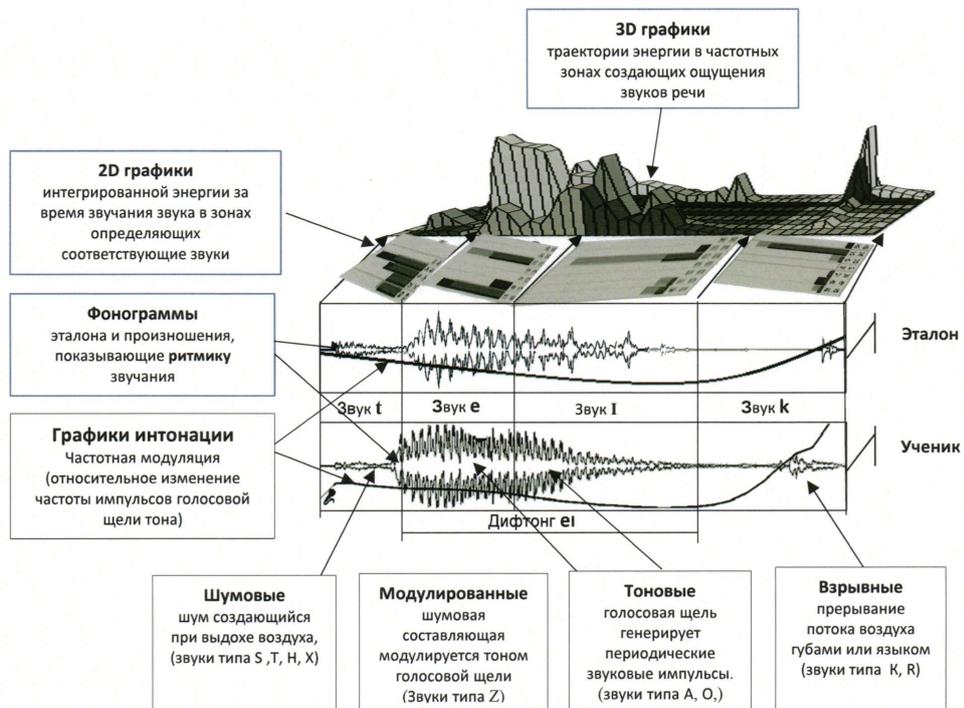
Фиг. 3

Звук [о]			Звук [у]			Звук [а]			Звук [и]			Звук [э]			Звук [ы]		
100-250	250-310	315-406	100-250	250-310	315-406	100-250	250-310	315-406	100-250	250-310	315-406	100-250	250-310	315-406	100-250	250-310	315-406
406-730	730-1130	1020-1230	406-730	730-1130	1020-1230	406-730	730-1130	1020-1230	406-730	730-1130	1020-1230	406-730	730-1130	1020-1230	406-730	730-1130	1020-1230
1230-1700	1700-2000	2200-3360	1230-1700	1700-2000	2200-3360	1230-1700	1700-2000	2200-3360	1230-1700	1700-2000	2200-3360	1230-1700	1700-2000	2200-3360	1230-1700	1700-2000	2200-3360
начало	конец	ширина															
200	400	400	300	700	400	200	1200	1000	-	-	-	300	700	500	200	500	300
-	-	-	-	-	-	-	-	-	1800	2300	300	1000	3300	3000	2000	3300	1300

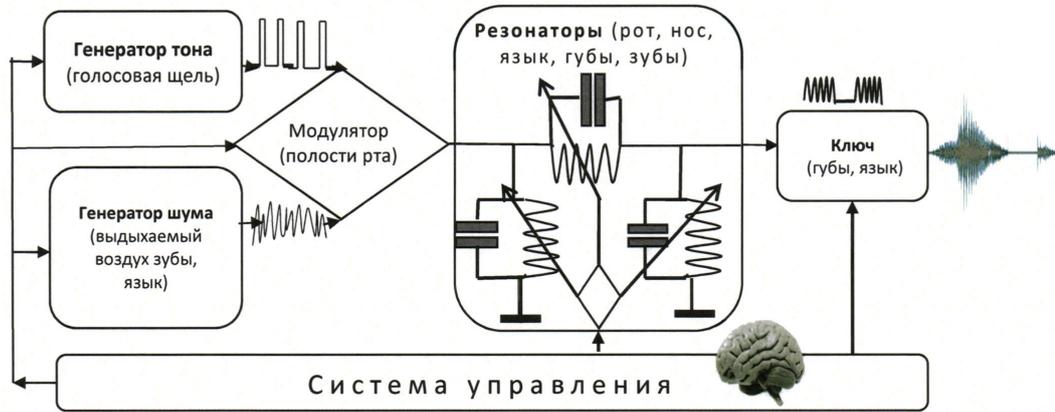
Фиг. 3а



Фиг. 3б



Фиг. 4



Фиг. 5