

УДК 004.934.2

Д. А. Ярославцева, А. Ю. Тычков

АНАЛИТИЧЕСКИЙ ОБЗОР СПОСОБОВ И АЛГОРИТМОВ ОЦЕНКИ СОСТОЯНИЯ ГОЛОСОВЫХ СКЛАДОК И ГОРТАНИ

D. A. Yaroslavtseva, A. Y. Tychkov

ANALYTICAL REVIEW OF THE METHODS AND ALGORITHMS OF ESTIMATION OF THE CONDITION OF THE VOCAL FOLDS AND LARYNX

А н н о т а ц и я. Представлены результаты аналитического обзора и критический анализ способов и алгоритмов оценки состояния голосовых складок и гортани. По результатам проведенного анализа предложен обобщенный алгоритм обнаружения и распознавания параметров речевых сигналов для оценки состояния голосовых складок и гортани.

A b s t r a c t. The article contains the results of the analytical overview and critical analysis of the methods and algorithms of estimation of the condition of the vocal folds, and larynx. The result of the analysis a generalized algorithm for detection and identification of the parameters of voice signals to assess the state of the vocal folds and larynx.

К л ю ч е в ы е с л о в а: речевой аппарат, голосовые складки, гортань, речевой сигнал, заболевания голосовых складок и гортани.

K e y w o r d s: speech apparatus, voice folds, larynx, the speech signal, diseases of the vocal folds and larynx.

Значительную роль в развитии теории и практики создания методов, способов и средств регистрации и обработки речевых сигналов сыграла необходимость у врачей оценивать информацию с помощью объективных количественных показателей с применением новых математических методов обработки.

В настоящее время значительно выросло количество специалистов голосоречевых профессий, увеличилась нагрузка на голосовой аппарат человека за счет появления различных средств коммуникации в условиях научно-технического прогресса. Это обусловило увеличение заболеваний голосового аппарата среди населения России и всего мира [1, 2]. По литературным данным, распространенность заболеваний гортани с каждым годом увеличивается [3]. Так, значительно увеличилась заболеваемость голосового аппарата среди учителей с 30–40 % в 50-е годы XX в. до 60–80 % в начале XXI в. [4, 5].

При перегрузке голосовых связок, хронических аллергических заболеваниях гортани, постоянном вдыхании раздражающих веществ, например, табачного дыма, в области голосовых связок может возникать постоянный отек слизистой оболочки, приводящий к охриплости и изменению тембра голоса. При таких симптомах немедленное обращение к отоларингологу необходимо для исключения онкологических заболеваний [6].

Последствия большинства заболеваний голосовых складок и гортани оказывают негативное влияние не только на речевую функцию пациента, но и вызывают тяжелое нарушение всех жизненно важных функций организма, а некоторые острые заболевания гортани могут в редких случаях привести к смерти больного от удушья.

При нарушениях голоса возникают социальные последствия в ограниченности общения вообще и в профессиональной деятельности в частности. Страх перед длительной нетрудо-

способностью, перед сменой профессии может приводить к снижению адаптации у взрослых, а у детей – к отставанию в физическом и умственном развитии.

Для полного восстановления голоса необходима диагностика речевого аппарата на ранних стадиях заболеваний, но большинство методов сложны в проведении и требуют высокой квалификации врача. Поэтому возникла необходимость в разработке систем и способов диагностики речевого аппарата по голосу говорящего с использованием современных достижений в области электроники и вычислительной техники.

Приведем критический анализ методов, способов и алгоритмов оценки состояния голосовых складок и гортани.

Информация о состоянии речевого аппарата может быть получена путем:

1) исследования состояния органов речеобразующего тракта врачом: осмотра, ощупывания (пальпация); простукивания (перкуссия); выслушивания (аускультация);

2) применения инвазивных (лабораторных) методов исследований: анализа крови и мочи; анализа биосубстрата; гормональных исследований;

3) применения неинвазивных (инструментальных) методов исследований: ларингоскопии, рентгеноскопии, томографии, бронхографии, эндоскопических методов исследования, виброметрии и др.

Совместное использование всех способов регистрации информации о состоянии речевого аппарата позволит эффективно выявлять заболевания органов речеобразующего тракта.

Одним из наиболее эффективных методов исследования фонаторной функции гортани является **виброметрия** [7].

К недостаткам относят сложную аппаратную часть, высокую квалификацию врача отоларинголога.

Для диагностики заболеваний двигательного аппарата гортани часто используют метод **реография гортани** [7].

Недостатком данного метода является сложность проведения обследования, для постановки диагноза необходима высокая квалификация врача.

Ларингоскопия (др.-греч. λάρυγξ – гортань + σκοπέω – наблюдаю, исследую) – метод визуального исследования гортани. Выделяют непрямую, прямую, ретроградную ларингоскопию. Данная методика выполняется с целью осмотра гортани при диагностических и лечебных мероприятиях. На рис. 1, 2 представлено схематическое изображение этапов введения ларингоскопа.

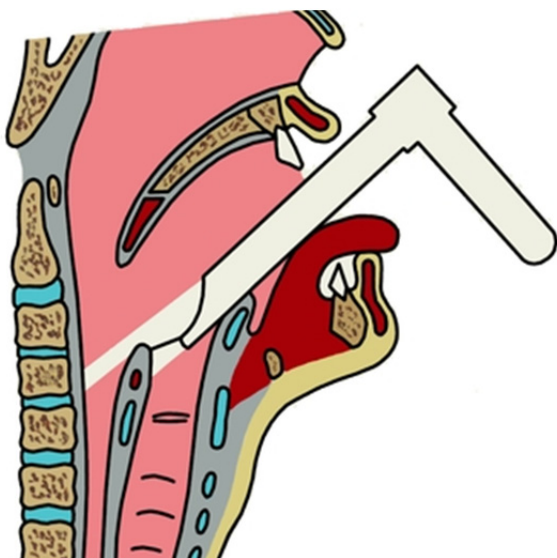


Рис. 1. Первый этап введения ларингоскопа

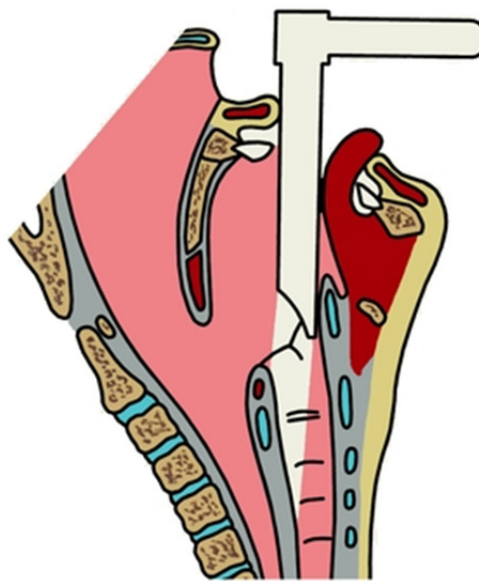


Рис. 2. Второй этап введения ларингоскопа

Недостатками являются значительное неудобство для пациента, необходимость дополнительной подсветки ларингоскопа, возможность повреждения гортани при вводе ларингоскопа.

Для исследования состояния гортани и голосовых складок используют **рентгеноскопию**. К недостаткам относят относительно высокую дозу облучения; низкое пространственное разрешение.

Исследование состояния органов дыхания специалистом обычно является субъективным методом, носит качественный характер и требует, как правило, высокой квалификации самого врача.

Требования стерильности, опасность заражения ВИЧ-инфекцией, дороговизна обслуживания ограничивают применение инвазивных методов исследования. Поэтому для контроля состояния органов дыхания целесообразно применение неинвазивных методов исследования.

В зависимости от способа получения информации о состоянии органов дыхания неинвазивные методы исследования подразделяются на четыре основных группы: лучевые методы; радионуклидные методы; биохимические и биофизические методы; нагрузочные методы и функциональные тесты.

Несмотря на качественный прорыв в области обследования неинвазивными методами состояния органов дыхания, существует ограничение в их использовании, связанное в основном с расшифровкой регистрируемой информации специалистом вручную. При этом скорость расшифровки остается относительно низкой, качество подвержено влиянию человеческого фактора.

Таким образом, необходимо улучшение работы методов, способов и систем регистрации и обработки речевых сигналов, которое может быть достигнуто за счет совершенствования и разработки новых методов, способов и систем обработки информации, обеспечивающих повышение точности автоматизированного обнаружения, распознавания информативных участков и измерения информативных параметров речевых сигналов. Все это позволит повысить доверие, как у пациента, так и у специалиста, к современным цифровым методам и системам обработки цифровых речевых сигналов и, соответственно, обеспечить эффективное принятие диагностических решений и повысить качество диагностики состояния органов дыхания.

Наиболее часто используемыми методами обработки речевых сигналов являются преобразование Фурье, вейвлет-преобразование, преобразование Гильберта–Хуанга, скрытые Марковские модели и др.

Известна система для анализа и формирования изображения шума дыхательных путей [8]. Данная система содержит многослойную матричную структуру микрофона, прикрепляемую к плоскости спины пациента, многоканальный аналого-цифровой преобразователь, память, процессор, предназначенный для обработки сигналов, хранящихся в памяти, монитор для отображения изображения полученных сигналов. Функциональная схема системы анализа и формирования изображения шумов дыхательных путей представлена на рис. 3.

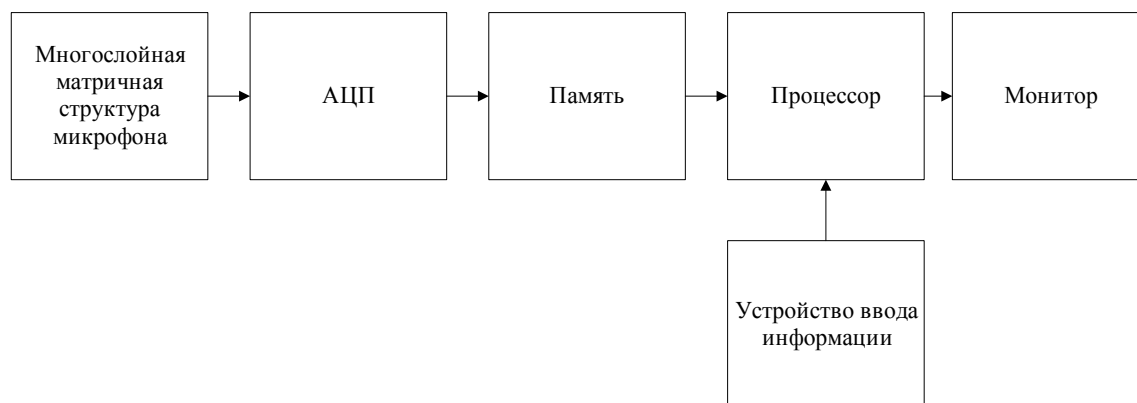


Рис. 3. Функциональная схема системы анализа и формирования изображения шумов дыхательных путей

Работа данной системы заключается в регистрации N сигналов шума дыхательных путей, определении на их основе средней акустической энергии, выборе интервала времени, определении средней акустической энергии на этом интервале и выводе ее в виде двумерного изображения. Алгоритм работы данной системы представлен на рис. 4.

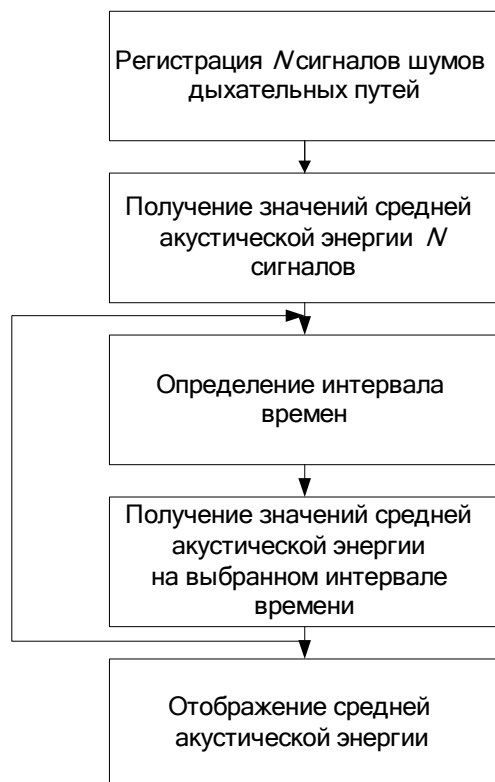


Рис. 4. Алгоритм работы системы анализа и формирования изображения шумов дыхательных путей

Недостатками данного способа являются влияние низкочастотных шумов и физиологических помех, связанных с дыханием и глотательным движением пациента, обусловленных регистрацией функциональной активности других (помимо легких) органов и систем организма: помех от движения; мышечных шумов; потенциалов, связанных с глотательными движениями. Другим недостатком данного изобретения является тактильный метод регистрации информации о состоянии легких и протекающих в них шумах, приводящий к повышенному дискомфорту пациента и соответственно снижению точности регистрируемой информации. Следующим недостатком данного изобретения является его сложная аппаратная реализация, связанная с применением многослойной матричной структуры микрофона и соответственно приводящая к дополнительным погрешностям при построении математической модели изображения шума дыхательных путей.

Функциональная схема устройства исследования функционального состояния голосовых складок [9] представлена на рис. 5.

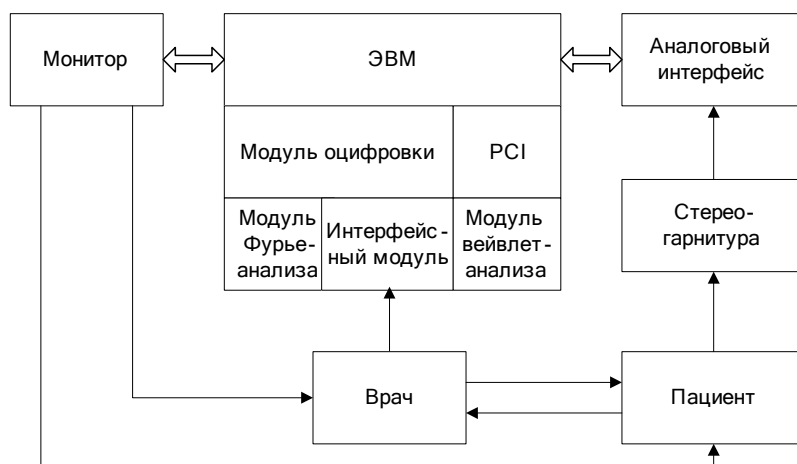


Рис. 5. Функциональная схема устройства исследования функционального состояния голосовых складок

Устройство исследования функционального состояния голосовых складок содержит стереогарнитуру, предназначенную для подключения к пациенту, аналоговый интерфейс, выходы которого подключены к выходу стереогарнитуры, ЭВМ, подключенный к выходу аналогового интерфейса, монитор, подключенный к ЭВМ и предназначенный для реализации взаимодействия между врачом, пациентом и модулем оцифровки сигнала ЭВМ.

Алгоритм работы устройства исследования функционального состояния голосовых складок: сначала происходит представление сигнала голоса в виде последовательности дискретных отсчетов, затем данную последовательность представляют в виде вейвлет-плоскости, полученную вейвлет-плоскость сегментируют и выбирают пять характерных признаков, значения которых рассчитаны для некоторых классов заболеваний голосовых складок. На рис. 6 приведен алгоритм работы известного способа исследования функционального состояния голосовых складок.

Недостатками способа исследования функционального состояния голосовых складок являются низкая специфичность, обусловленная применением вейвлет-преобразования к речевому сигналу, а также низкая эффективность, обусловленная **невозможностью** адаптации базового вейвлета к локальным особенностям зарегистрированного сигнала.

Следующим недостатком известного способа является то, что значения характерных признаков, относящиеся к классам ряда заболеваний органов речеобразующего тракта, коррелируют между собой.

Обзор существующих способов и алгоритмов показал, что ни один из них не позволяет с необходимой точностью провести исследование функционального состояния голосовых складок и гортани.



Рис. 6. Алгоритм работы способа исследования функционального состояния голосовых складок

Для разработки эффективной методики исследования заболеваний органов дыхания по речи говорящего необходимо разработать базу данных [10–12] контрольных слов, фраз и выражений, позволяющую комплексно оценить работу внутренних органов дыхания. Предложенная база имеет три составляющих блока, представленных в виде таблиц реляционной СУБД Microsoft OfficeAccess.

Первый блок БД сформирован путем проведения аналитического обзора и анализа информации о строении звуков речи и проблем, возникающих при их произношении. Данный блок содержит «проблемные» звуки, неправильное произношение которых позволяет комплексно оценить работу внутренних органов дыхания. На основе представленных звуков (43 звука в русском языке) составлены слова и предложения (табл. 1), позволяющие специалисту определять состояние работы внутренних органов дыхания и сформировать предварительное заключение.

Таблица 1

Первый блок БД

а) Проблемные звуки

«Проблемный звук»	Слово	Классификация звуков, составляющих слово	Предложения, содержащие это слово	Причины неправильного произношения	Диагноз
С, З, Ц	Сцена	2 тональных, 1 сонорный, 2 свистящих (1 аффрикат)	Цицерон тридцать раз облизнулся, заметив бесценное заземление на сцене	Дефекты свистящих звуков	Шепелявость Сигматизм губно-зубной Сигматизм призубной Сигматизм шипящий

Первый столбец первого блока базы данных содержит «проблемные звуки», например «с, з, ц», из которых составляются слова (в данном случае слово «сцена», которое записывается во второй столбец). Третий столбец строки несет информацию о фонетическом составе слова, к примеру, слово «сцена» состоит из 2 тональных, 1 сонорного, 2 свистящих (1 аффрикат) звуков. В четвертом столбце приведены предложения, структурные единицы которых состоят преимущественно из «проблемных звуков», эти предложения включают слова из второго столбца, например: «Цицерон тридцать раз облизнулся, заметив бесценное заземление на сцене». Пятый столбец отображает причины неправильного произношения слов и предложений. Шестой столбец позволяет сформировать предварительный диагноз, который поможет поставить заключение специалисту.

Второй блок базы данных является базисом для итогового третьего (табл. 2). Содержание второго блока основано на анализе информации об артикуляции и имеет связь между всеми звуками русского языка и органами речеобразования. Первый столбец содержит все звуки русского языка. Все органы дыхания разделены на две группы: активно и пассивно участвующие в образовании данного звука. В последнем случае это, как правило, органы дыхания. Например, в образовании звука «э» (первый столбец) активно участвуют кончик языка и гортань (второй столбец), пассивное участие принимают легкие, грудная клетка, диафрагма, нижние зубы (третий столбец).

Таблица 2

Второй блок БД

б) Звуки – органы

Звуки	Органы, активно участвующие в образовании данного звука	Органы, пассивно участвующие в образовании данного звука
а	Гортань	Легкие, грудная клетка, диафрагма
у	Губы, средняя и задняя части и кончик языка, гортань	Легкие, нижние зубы, грудная клетка, диафрагма
ы	Средняя и задняя части спинки языка, гортань	Легкие, грудная клетка, диафрагма
э	Кончик языка, гортань	Легкие, грудная клетка, диафрагма, нижние зубы

Третий блок базы данных предназначен для работы специалиста (табл. 3). Первый ее столбец содержит диагностируемые органы речеобразования, второй – звуки, в образовании которых эти органы участвуют; содержимое этих столбцов является результатом обработки данных второй таблицы. Получается, что по произношению предложений возможно определить состояние органов речеобразования. Однако это не однозначно верно, так как одни и те

же звуки зачастую соответствуют нескольким органам, и одно предложение может диагностировать группу органов. Для этого были введены четвертый и пятый столбцы, разделяющие звуки по степени участия в их образовании соответствующих органов на основные и дополнительные соответственно. Из основных звуков составляются слова, которые могут конкретизировать диагностируемый орган, и записываются в шестой столбец. При составлении уделялось внимание дополнительным звукам, так как повторяются и основные, но при этом дополнительные могут различаться. Фонации учитываются при подборе слов на основе сонорных и тональных звуков. Для конкретизации гортани используются слова: аорта, ананас, бархат, баян, май – гласные в них наиболее вокализованы, тогда как слова, конкретизирующие полость носа, также имеют тональные звуки между сонорными, но «заглатываются» при произношении (намыленный, минный, нанизан).

Таблица 3

Третий блок БД

в) Органы – звуки – слова – предложения

Орган	Звуки речи	Предложения	Основные	Дополнительные	Слова
Голосовые связки	б, б', в, в', г, г', д, д', ж, з, з', й, л, л', м, м', н, н', р, р', а, и, о, э, ы, у	Лара врала, что была королевой Урала. Лор просил, чтобы я поорал	л, л', м, м', н, н', р, р', а, и, о, э, ы, у	б, б', в, в', г, г', д, д', ж, з, з', й	Нормально, Эллина, радио, вандал
Гортань	а, и, о, э, ы	Элементарные мысли об ананасах, бананах или кокосовом соке – это идиллия	а, и, о, э, ы		Аорта, ананас, бархат, баян, май

Эффективная работа системы контроля и оценки состояния здоровья органов внутреннего дыхания может быть обеспечена только при условии, что меры по обработке и анализу речевых сигналов предусматриваются на всех этапах преобразования информации.

Для решения поставленной задачи предлагается обобщенный алгоритм обнаружения параметров речевых сигналов при заболеваниях голосовых складок и гортани (рис. 7).

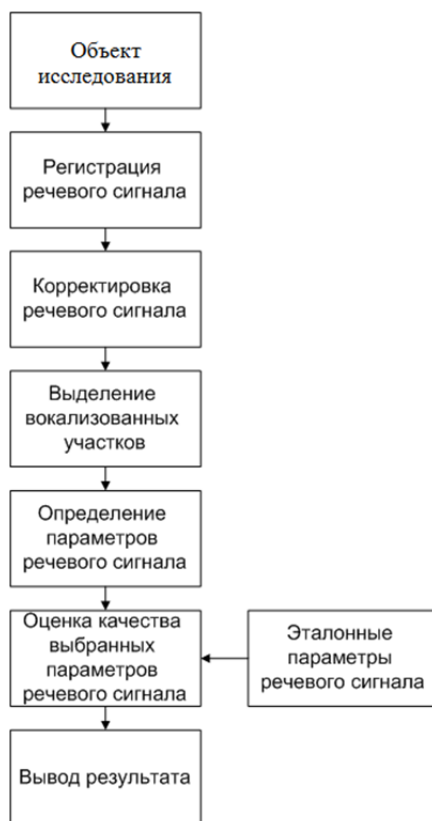


Рис. 7. Обобщенный алгоритм анализа параметров речевых сигналов при заболеваниях голосовых складок и гортани

Новый способ обработки и анализа речевых сигналов должен основываться на сочетании различных алгоритмов. В основе разработанного способа на этапах обработки и анализа речевых сигналов предложено использовать преобразование Гильберта–Хуанга и разложение сигнала на эмпирические моды. Эти операции, по мнению авторов предлагаемого способа, позволяют увеличить точность разграничения параметров речевого сигнала для определения заболеваний голосовых складок и гортани.

На этапах обработки и анализа речевых сигналов необходимо разработать алгоритмы, методику и систему критериев, позволяющих повысить качество регистрируемого речевого сигнала, эффективность обнаружения и определения информативных параметров сигнала на основе разработанной базы данных речевых слов, фраз и выражений.

В предложенном алгоритме выполняется последовательность действий:

1. Аудиозахват речевого сигнала – регистрация речевого сигнала с помощью сертифицированных приборов, баз данных (БД) АРМ врача и БД Internet.
2. Корректировка речевого сигнала – удаление пауз в начале и конце зарегистрированного сигнала.
3. Выделение из речевого сигнала всех видов фонем – сегментация на фонемы.
4. Анализ спектральных и статистических характеристик фонем – определение информативных параметров фонем.
5. Сравнение полученных характеристик с эталонами – оценка качества фонем.

Чтобы разработать алгоритм, позволяющий оценивать состояние голосовых складок и гортани по голосу говорящего, необходимо понимать процессы, которые происходят в речевом тракте, и знать свойства речевого сигнала.

Речевой сигнал имеет двойственную природу – с одной стороны, это обычный акустический сигнал, который представляет собой процесс распространения энергии акустических колебаний в упругой среде. Как любой акустический сигнал, он может быть представлен в виде звуковых волн, представляющих собой распространение процессов сжатия и разрежения частиц среды, формы фронтов которых зависят от свойств источника и условий распространения. Поэтому, как и другие акустические сигналы, речь характеризуется определенным набором объективных характеристик: зависимостью звукового давления от времени (временной структурой звуковой волны), длительностью звучания, спектральным составом, местом расположения источника в пространстве [13].

С другой стороны, речь как физическое явление вызывает определенные субъективные слуховые ощущения (громкости, высоты, тембра, локализации, маскировки и др.).

Человеческий голос, его акустические свойства, механизмы его порождения изучают самые различные науки: физиология, фонетика, фониатрия, логопедия и др. Поскольку голосовой феномен – явление не только физиологическое, но и физическое, он становится предметом изучения такого раздела физики, как акустика, которая дает четкие характеристики каждого воспроизведенного звука. Согласно акустике звуком считается распространение колебаний в упругой среде. Человек и говорит, и поет в воздушной среде, поэтому звук голоса – это колебание частиц воздуха, распространяющихся в виде волн сгущения и разрежения, как волны на воде, со скоростью 340 м/с при температуре +18°C [14].

Разработан обобщенный алгоритм обнаружения параметров речевых сигналов при заболеваниях голосовых складок и гортани, реализуемый с использованием разработанной верифицированной БД и преобразования Гильберта–Хуанга.

Список литературы

1. Аникеева, З. И. Клинические особенности нарушений голоса у населения мегаполиса / З. И. Аникеева // Вестник оториноларингологии. – 2007. – № 1. – С. 14–21.
2. Дайхес, Н. А. Клиническая патология гортани / Н. А. Дайхес, В. П. Быкова. – М. : Медицинское информационное агентство, 2009. – 160 с.
3. Крюков, А. И. Показатели заболеваемости и качество оказания амбулаторной ЛОР-помощи больным с патологией уха и верхних дыхательных путей в городе Москве / А. И. Крюков, Р. Б. Хамзалиева, А. Ф. Захарова // Российская оториноларингология. – 2008. – № 3. – С. 23–26.
4. Chevalier, D. Acute disorder of speech / D. Chevalier, G. Mortuaire // Dysphonia. Rev Prat. – 2005. – № 11. – P. 1217–1221.

5. McAllister, A. Children's voice and voice disorders / A. McAllister, P. Sjölander // *Semin Speech Lang.* – 2013. – № 34 (2). – P. 71–79.
6. URL: <http://lorclinics.ru>
7. Бабияк, В. И. Оториноларингология : руководство / В. И. Бабияк, М. И. Говорун, Я. А. Накатис. – СПб. : Питер, 2009. – Т. 2. – С. 56–63.
8. Пат. 2004124247/14 Российская Федерация. Система для анализа и формирования изображения шума дыхательных путей / Кушнир И., Ботбол М. – Оpubл. 20.01.2008, Бюл. № 2.
9. Пат. 2006116845/14 Российская Федерация. Способ исследования функционального состояния голосовых складок / Плешков И. В., Филист С. А., Краснова Н. В., Блинков Ю. А. – Оpubл. 27.12.2007, Бюл. № 36.
10. Алимуратов, А. К. Фильтрация речевых сигналов с использованием метода множественной декомпозиции и оценки энергии эмпирических мод / А. К. Алимуратов, П. П. Чураков, А. Ю. Тычков // *Известия высших учебных заведений. Поволжский регион. Технические науки.* – 2012. – № 4. – С. 50–61.
11. Исследование и разработка верифицированной базы данных звуков речи для анализа речевых сигналов / А. К. Алимуратов, А. Ю. Тычков, Ю. С. Квитка, Д. А. Ярославцева // *Современные проблемы компьютерных наук (СПКН-2013) : сб. материалов I Междунар. науч.-практ. конф., посвящ. 70-летию образования Пензенского государственного университета / под ред. Л. Р. Фионовой.* – Пенза : Изд-во ПГУ, 2013. – С. 156–159.
12. Свидетельство 20136621400 Российская Федерация. Верифицированная база речевых сигналов / А. Ю. Тычков, А. К. Алимуратов, Ю. С. Квитка, Д. А. Ярославцева. – Оpubл. 06.11.2013.
13. Алдошина, И. Акустические характеристики речи / И. Алдошина // *Звукорежиссер.* – 2002. – Вып. 4. – С. 33–36.
14. Лаврова, Е. В. Логопедия / Е. В. Лаврова // *Основы фонопедии.* – М., 2007. – С. 12–15.

Ярославцева Дарья Александровна

студентка,
Пензенский государственный университет
E-mail: levik_92@mail.ru

Yaroslavtseva Dar'ya Aleksandrovna

student,
Penza State University

Тычков Александр Юрьевич

кандидат технических наук,
директор студенческого научно-производственного
бизнес-инкубатора,
Пензенский государственный университет
E-mail: tychkov_a@pnzgu.ru

Tychkov Aleksander Yur'evich

candidate of technical sciences,
director of student research
and production business incubator,
Penza State University

УДК 004.934.2

Ярославцева, Д. А.

Аналитический обзор способов и алгоритмов оценки состояния голосовых складок и гортани / Д. А. Ярославцева, А. Ю. Тычков // *Измерение. Мониторинг. Управление. Контроль.* – 2014. – № 2 (8). – С. 72–80.