

ОТЗЫВ

официального оппонента д.т.н. Лобанова Бориса Мефодьевича

на диссертацию Якимука Алексея Юрьевича «Алгоритмы анализа частоты основного тона вокального исполнения», представленную на соискание ученой степени кандидата технических наук по специальности 05.13.17 – «Теоретические основы информатики»

1 Актуальность работы

Диссертационная работа направлена на решение научно-технических задач анализа речевого сигнала, а именно, задач распознавания звучащих нот в вокальном исполнении и создания на этой основе программных средств обучения пению.

Развитие сферы речевых технологий привело к внедрению программных средств, распознающих или синтезирующих человеческую речь во многие сферы жизни, в том числе, для обучения речевых навыков. Стоит отметить, что применение программ для развития речевых навыков (таких как обучение произношению на иностранном языке или искусству пения заданной последовательности нот) накладывает на них требование к высокой точности при оценке корректности выполнения упражнений. Проведенное автором анализ источников показал, что существующая форма обучения вокалу подразумевает взаимодействие с репетитором. Самостоятельное выполнение упражнений редко способно развить музыкальный слух, т.к. отсутствует обратная связь, контролирующая динамику изменения фундаментальной частоты - $F_0(t)$ - в вокальном исполнении. Автор отмечает, что существующие алгоритмы не позволяют вычислить значения $F_0(t)$ с высокой точностью из-за наличия высокого процента грубых ошибок и узкого спектра охватываемых частот.

В качестве цели работы Якимука А.Ю. значится повышение качества распознавания звучащих нот в вокальном исполнении. Направленность на решение этой задачи, имеющей существенное значение в развитии речевых технологий, позволяет классифицировать тематику рассматриваемой диссертации как актуальную.

2 Структура и содержание диссертации

Диссертационная работа Якимука А.Ю. содержит введение, четыре главы, заключение, три приложения и список источников из 157 наименований. Объем диссертационной работы составляет 121 страницу, включая 12 таблиц и 53 рисунка.

Оформление и структура работы в целом соответствует требованиям, предъявляемым к кандидатским диссертациям. Содержание автореферата соответствует основным идеям, результатам, выводам и положениям диссертации.

Во введении автор обосновывает актуальность темы исследования и формулирует цель работы. Далее автор излагает основные результаты проведенных исследований и показывает их научную новизну, теоретическую и практическую значимость, отражает основные положения, выносимые на защиту.

В первой главе автором приводятся результаты проведенного исследования. Кратко описаны алгоритмы анализа частоты основного тона и сегментации, а также примеры их применения в задачах, близких к анализу вокальных исполнений. Приведены показатели рассмотренных алгоритмов с оценкой применимости в задаче распознавания нот в

вокальном исполнении. Приведено сравнение программ-аналогов с указанием их особенностей и краткий обзор исследований по применению программных средств в сфере обучения пению.

Во второй главе приведено описание процесса формирования наборов шаблонов для определения частот основного тона в вокальном исполнении. Описана модификация выбранной модели слуховой системы человека и приводятся результаты проведенного тестирования работы алгоритма идентификации частот на сгенерированных синусоидальных сигналах.

В третьей главе описывается разработанный алгоритм распознавания нот в вокальном исполнении, включающий в себя этапы сегментации и идентификации нот. Приводится обоснование выбора метода вычисления границ звучания ноты. Предложены стратегии, по которым проведен сбор аудиозаписей с пением. Показаны результаты тестирования разработанных алгоритмов на собранных аудиозаписях.

В четвертой главе описывается разработанный автором программный комплекс. Приводятся результаты тестирования на аудиозаписях с различными подходами к вокальному исполнению.

В заключении приведены основные выводы и результаты диссертационного исследования.

3 Новизна полученных результатов

Результатами диссертации, обладающими признаками научной новизны, являются:

1) Модификация известной модели слуховой системы человека, описанной в диссертации А. А. Конева, позволившая расширить диапазон частот в 2 раза по сравнению с исходной моделью и отличающаяся возможностью произвольного указания частотных границ определения тона.

2) Алгоритм создания шаблонов для обнаружения частоты основного тона, отличающийся возможностью автоматической генерации наборов шаблонов с произвольным заданием граничных частот определения основного тона.

3) Алгоритм распознавания нот с минимальной длительностью звучания, обусловленной особенностями слуховой системы человека.

4) Алгоритмы сегментации аудио сигнала с вокальным исполнением на вокализованные и невокализованные участки

4 Практическая и теоретическая ценность и внедрение результатов

Практическая значимость полученных автором результатов заключается в их использовании для решения практических задач:

- автоматического определения нот в вокальном исполнении;
- создание акустической базы данных, в том числе, женских и мужских певческих голосов, для тестирования алгоритмов и программ автоматического определения нот.

Результаты внедрены в деятельность «Элекард-ЦТП» в рамках проекта по дистанционному обучению вокалу в формате видеоконференций. Значимость результатов работы подтверждена актом внедрения в деятельность «Элекард-ЦТП»: внедрение позволило снизить объем трафика, передаваемого по сети, более чем на 90% за счет

перехода от передачи аудиозаписей с пением в формате wav к отправлению преподавателю текстового сообщения с abc-нотацией распознанных нот.

Теоретическая значимость результатов диссертационного исследования заключается в развитии методов анализа частоты основного тона речевого сигнала. Модификация математической модели слуховой системы человека позволила расширить диапазон определения частот основного тона до 800 Гц. Теоретические результаты используются в учебном процессе Томского государственного университета систем управления и радиоэлектроники, что подтверждается актом внедрения.

5 Достоверность и обоснованность основных результатов и выводов

Достоверность и обоснованность результатов и выводов подтверждается результатами проведенных численных экспериментов с использованием реальных данных, а также путём сопоставления результатов, полученных в диссертации, с результатами экспертных оценок.

Достоверность разработанной акустической базы данных женских и мужских певческих голосов подтверждена успешным тестированием её в составе программной системы “**Singing Lessons - IntonTrainer**” (см. сайт: <https://intontrainer.by/>).

6 Рекомендации по использованию результатов работы

Результаты диссертационного исследования рекомендуются к применению на практике в обучении вокалу в музыкальных школах.

В качестве рекомендации для дальнейшего развития исследований можно предложить расширение авторского подхода, направленное на распознавание нот в вокальном исполнении с резким изменением частот основного тона, наблюдаемом при пении с такими подходами как вибрато и глиссандо. Кроме того, может быть предложено направление на реализацию данных алгоритмов в программном комплексе в режиме реального времени с включением набора конкретных упражнений, что позволит внедрить программу в музыкальные школы.

7 Публикации по теме диссертации

Результаты диссертации опубликованы в 19 научных работах, в том числе 3 публикации в рецензируемых журналах из перечня ВАК по специальности 05.13.17 – «Теоретические основы информатики» и 2 публикации в рецензируемых журналах из перечня ВАК по другим специальностям, 2 публикации в научных изданиях, индексируемых Scopus, 14 публикаций в тезисах и материалах научных конференций, получены 2 свидетельства о государственной регистрации программы для ЭВМ.

Опубликованные материалы достаточно полно отражают полученные в диссертации теоретические и практические результаты.

Наиболее значимые публикации соискателя по теме диссертации:

1. **Якимук А.Ю.**, Конев А.А. Алгоритм сегментации речевого сигнала на основе значений минимальной меры различия // Информатика и системы управления. – 2018. – № 2 (56). – С. 108-121.

2. Катаева Е.С., Якимук А.Ю. Применение метода выделения синхронности при оценке сходства вокальных исполнений // Доклады Томского государственного университета систем управления и радиоэлектроники. – 2019. – Т. 22. – №3.

3. Якимук А.Ю., Конев А.А., Осипов А.О. Программный комплекс для автоматизации моделирования сегментации речевых сигналов и вокальных исполнений // Вестник Иркутского государственного технического университета. – 2017. – Т. 21. – № 10 (129). – С. 53-64.

8 Замечания по диссертации

1. Отсутствует перечень условных обозначений, что существенно затрудняет восприятие аббревиатур и формул, используемых в тексте диссертации.
2. Диссертационная работа, по утверждению её автора, основана на математической модели слуховой системы человека. Однако, в диссертации полностью отсутствует раздел, посвящённый обзору многочисленных современных моделей слуховой системы человека. Приводится лишь единственная ссылка на кандидатскую диссертацию Конева А. А. «Модель и алгоритмы анализа и сегментации речевого сигнала» без описания, однако, существа приведенной там модели.
3. В первой главе (стр. 21) автор утверждает, что *«Проведенный анализ алгоритмов анализа частоты основного тона (таблица 1.2) показал, что существующие алгоритмы в текущем виде не могут быть применены в задаче обработки вокальных исполнений»*. Соответственно, в выводах по 1-й главе заявлено: *«было определено, что наименьшим процентом ошибок при идентификации частот основного тона обладает алгоритм, основанный на математической модели слуховой системы человека»*. Однако, из текста главы остаётся не ясным: с использованием какой методики и на каком тестовом материале это определено?
4. В выводах по 3-й главе указывается, что *«данный подход был также исследован при анализе шепотной речи»* (стр. 75), однако, экспериментального подтверждения того, как это возможно при пении, отсутствует. И, вообще, *«шепотное пение?»* - крайне не ясное понятие.
5. В 4-й главе «Разработка программного комплекса исследования вокализованной речи» лишь кратко описаны структура комплекса и его функционирование. Отсутствует, однако, описание интерфейса комплекса, не приводится также руководство пользователя. Их было бы желательно дать в качестве приложений к диссертации.

Приведенные замечания не снижают общей положительной оценки диссертационной работы и не ставят под сомнение значимость полученных результатов.

9 Соответствие темы диссертации заявленной научной специальности

Тема и положения, выносимые на защиту, соответствуют паспорту специальности 05.13.17 – «Теоретические основы информатики» по пункту 5:

«Разработка и исследование моделей и алгоритмов анализа данных, обнаружения закономерностей в данных и их извлечениях. разработка и исследование методов и алгоритмов анализа текста, устной речи и изображений».

10 Оценка диссертации

Диссертация Якимука А.Ю. является самостоятельной научно-квалификационной работой, в которой даны научно обоснованные решения по повышению качества распознавания звучащих нот в вокальном исполнении, разработке модифицированной модели анализа акустических сигналов пения человеком и её использование в компьютеризированной системе обучения.

Диссертация отвечает требованиям пункта 9 «Положения о порядке присуждении ученых степеней» ВАК РФ по актуальности, научной новизне, значимости, опубликованности и апробированности, а ее автор, Якимук Алексей Юрьевич, достоин присуждения ему ученой степени кандидата технических наук по специальности 05.13.17 – «Теоретические основы информатики» (технические науки).

Официальный оппонент

доктор технических наук (05.13.01), главный научный сотрудник лаборатории распознавания и синтеза речи Государственного научного учреждения «Объединённый институт проблем информатики Национальной академии наук Беларуси»



Лобанов Борис Мефодьевич

Беларусь, 220012, г. Минск, ул. Сурганова, 6
Телефон: +375(17) 284 27 73
E-mail: lobanov@newman.bas-net.by

Подпись Лобанова Бориса Мефодьевича удостоверяю:

Учёный секретарь

Объединённого института проблем информатики НАН Беларуси



Горох Олег Владимирович

6.12.19