

УДК 004.934.2

Н.С. Репьюк, А.А. Конев

## Программный комплекс для исследования речевых сигналов

Представлен программный комплекс, предназначенный для исследования особенностей параметров речевого сигнала. Данный комплекс планируется применить для разработки параметрического описания звуков и(или) групп звуков. Представленный комплекс создавался на основе алгоритмов, разработанных А.А. Коневым с учетом предъявленных требований для модификации предыдущей версии комплекса. Описаны структура, архитектура и основные блоки данного комплекса.

**Ключевые слова:** программный комплекс, речевой сигнал, фильтрация, сегментация.

**DOI:** 10.21293/1818-0442-2025-28-1-93-99

Речевые технологии распознают, анализируют и синтезируют человеческий голос. Имитация речи, восприятие смысла фраз, преобразование речи в текст, работа с голосом как биометрической характеристикой – все это разные виды речевых технологий. Распознавание речи – наиболее перспективное направление исследований, поскольку речь является естественным способом коммуникации. Распознавание слитной речи и идентификация человека по голосу – особенно сложные проблемы для открытого множества дикторов.

По данным Fortune Business Insights [1], в 2023 г. мировой рынок распознавания речи составил 12,62 млрд долл., по прогнозам, в 2024 г. – 15,45 млрд долл.

Существует множество программных комплексов для распознавания речи для различных языков. Комплексы бывают для распознавания отдельных слов, словосочетаний, слитной и спонтанной речи. При этом алгоритмы работы в этих комплексах чаще всего строятся на скрытых Марковских моделях или на различных нейронных сетях, или на их гибриде [2]. Современные системы для распознавания речи требуют больших наборов записей различных людей и условий [3], чтобы обучить систему и избавиться от такого недостатка, как дикторозависимость.

Для решения задач в области речевых технологий требуется детальный анализ структуры речевого сигнала. При этом важно, что он должен проводиться с учетом особенностей слухового восприятия человека. Поэтому возникает потребность в программном обеспечении, с помощью которого можно исследовать особенности параметров речевого сигнала, чтобы в дальнейшем разработать параметрическое описание звуков и(или) групп звуков.

Целью данной работы является повышение эффективности проведения исследований при помощи разработанного программного комплекса. Эффективная работа пользователя с интерфейсом разработанного программного комплекса позволяет повысить качество проведения исследований речевых сигналов. Модернизация комплекса позволяет ускорить работу существующего комплекса и улучшает интеграцию различных алгоритмов и функций с комплексом.

### Обзор комплексов для изучения речевого сигнала

Создание параметрического описания звуков и(или) групп звуков требует подробного изучения параметров речи. Для изучения речевого сигнала существуют различные программные инструменты.

По запросу «acoustic analysis program» в исследовательской базе данных Mendeley была найдена 361 статья в различных журналах за последние три года (с 2022 по 2024 г.).

По запросу «acoustic analysis software» было найдено 628 статей за этот же период. При этом было найдено 493 статьи о применении комплекса Praat и 105 статей о применении Multi-Dimensional Voice Program (MDVP).

Еще одним из таких инструментов является программа для акустического анализа речевых сигналов WinPitchPro [4–6]. Данная программа работает не только с аудиофайлами различных форматов, но и с речевым сигналом в реальном времени.

MDVP [7–12] анализирует несколько параметров, таких как частота основного тона (ЧОТ), дрожание, мерцание, отношение шума к гармонике, индекс турбулентности голоса и индекс интенсивности амплитудного тремора.

Программное обеспечение Praat [7–13] используется для визуального представления речевого сигнала с помощью спектрограмм. Praat позволяет выполнять спектральный анализ, анализ высоты тона, формант и интенсивности, а также изучать дрожание, мерцание и голосовые разрывы. В данной программе есть функция автоматического распознавания текста по речевому сигналу и соотнесения текста и сигнала на графике.

Центр речевых технологий в России разрабатывает и производит программно-аппаратные и программные продукты для обработки речи и ее анализа. Один из таких продуктов – «ИКАР Лаб» [14]. Аппаратно-программный комплекс ИКАР Лаб 3 предназначен для проведения криминалистической экспертизы аудиозаписей и используется во внутренних силовых структурах.

Ранее в рамках диссертационной работы на соискание степени кандидата технических наук А.А. Коневым был разработан программный комплекс для изучения параметров речевого сигнала на языке Delphi [15, 16].

Ниже приведена сравнительная таблица программных средств для акустического анализа речевых сигналов.

Обзор и анализ рассмотренных программ для акустического анализа речевого сигнала выявил их главный недостаток, который не позволяет использо-

вать известные программные решения для исследования речевого сигнала, а именно, точность вычисления параметра ЧОТ. Комплекс А.А. Конева был взят

за основу для созданного программного комплекса, так как он имеет наименьшую погрешность вычисления ЧОТ.

#### Программные комплексы для анализа речевых сигналов

| Программный комплекс   | Работа в реальном времени  | Функциональные возможности   | Параметры  | Определение ЧОТ   | Формат построения графиков   | Выделение отдельных сегментов              | Источник |
|--|--|--|--|---|--|--|----------|
| WinPitchPro  | ПО позволяет в режиме реального времени анализировать и отображать просодические кривые (F0 и интенсивность) вместе с соответствующей спектрограммой | Для соответствующего сегмента можно ввести соответствующий символ из Unicode   | ЧОТ, частота гармоник, интенсивность сигнала   | AMDF, спектральная гребенка, спектральная кисть, автокорреляция и гребенка выбранных гармоник (5 методов). Процент грубых ошибок: 32,92% при частотах, близких к 500 Гц | Спектрографический анализ  | Присутствует выделение слов, слогов, фонем | [4–6]    |
| Praat  | Нет  | Для выделенного сегмента можно ввести метку с буквами, слогами и т.д. Есть автоматическое транскрибирование речи в текст | ЧОТ (F0), дрожание, мерцание, NHR (отношение шума к гармонике), VPI (индекс турбулентности голоса) и ATRI (индекс интенсивности амплитудного тренда) | Алгоритм, основанный на методе автокорреляции. Процент грубых ошибок: 3,96%   | Спектрографический анализ, формантный анализ, анализ интенсивности | Присутствует выделение слов, слогов, фонем | [7–12]   |
| MDVP   | Нет  | Для выделенного сегмента можно ввести метку с буквами, слогами и т.д.  | ЧОТ (F0), дрожание, мерцание, NHR, VPI и ATRI  | Алгоритм, основанный на методе автокорреляции. Процент грубых ошибок: 3,96%   | Спектрографический анализ, формантный анализ, анализ интенсивности | Присутствует выделение слов, слогов, фонем | [7–13]   |
| ИКАР ЛабЗ  | Нет  | Установление дословного содержания фонограммы с разделением по дикторам, идентификация диктора по фонограмме устной речи | ЧОТ  | Спектральный анализ сигнала фонограммы. Процент грубых ошибок: 32,92% при частотах, близких к 500 Гц  | Вейвлет-сонограмма, спектрографический анализ                      | Разделение графиков на сегменты            | [14]     |
| Основанный на математической модели слуховой системы человека (А.А. Конев) | Нет  | Построение графиков ЧОТ, частот гармоник и интенсивности сигналов от временных отчетов                                   | ЧОТ, частота гармоник, интенсивность сигнала   | Погрешность определения ЧОТ составляет не более 0,6%  | На координатной плоскости  | Нет  | [15, 16] |

Еще одним недостатком всех рассмотренных систем является то, что рассмотренные комплексы не транскрибируют русскую речь и нет возможности указания меток на русском языке для выделенных сегментов

#### Требования по модернизации предыдущей версии комплекса

Комплекс А.А. Конева был взят за основу для создания программного комплекса, однако у ком-

плекса А.А. Конева есть существенные недостатки: система является монолитной, что затрудняет внедрение новых модулей или доработки уже имеющихся, графический интерфейс не включает в себя пофонемное выделение участков на графике, что затрудняет процесс исследования параметров отдельных фонем. Программа разделена на три модуля, и каждый модуль имеет свое отдельное окно в программе, что затрудняет работу с комплексом.

Для улучшения работы комплекса было решено его модифицировать. Для этого были выявлены следующие требования для правильной работы программы:

1. *Язык программирования.* Программа должна работать с аудиофайлами и некоторым набором фильтров, поэтому комплекс должен обрабатывать большой объем данных. Для оптимизированной работы с памятью и многопоточностью был выбран язык программирования C#.

2. *Структура программного обеспечения.* Структура комплекса должна быть построена таким образом, чтобы была возможность разрабатывать и внедрять новые модули в систему. Помимо этого, должна быть возможность модификации уже имеющихся модулей комплекса. Для внедрения новых модулей в программу была выбрана модульная структура комплекса. При этом модули системы состоят из нескольких классов.

3. *Графический интерфейс.* Для удобного взаимодействия пользователя с системой необходимо разработать интуитивно понятный графический интерфейс. Интерфейс не должен состоять из множества окон, в которых предлагается выбор соответствующего файла.

4. *Работа с графиками.* В предыдущей версии комплекса по итоговым графикам не определяются фонемы и нельзя выделить изучаемый участок графика, что затрудняет анализ полученного результата. Для разметки графиков необходимо разработать ручное выделение отдельных фрагментов графиков и присвоение им имен. Программа должна иметь возможность ручного выделения участка по временным отчетам для детального изучения отдельных частей аудиозаписи.

5. *Фреймворк.* В качестве среды разработки была выбрана Visual Studio 2022. Для разработки десктопных приложений используется интерфейс программирования Windows Forms. Но данный интерфейс не позволяет разрабатывать кроссплатформенные приложения. Поэтому для кроссплатформенности готового программного комплекса необходимо выбрать удобный фреймворк с множеством открытых библиотек.

#### **Описание архитектуры программного комплекса**

Согласно выделенным требованиям, было необходимо модифицировать программный комплекс для изучения речевого сигнала. Для оптимизации работы программного комплекса с большим объемом данных язык программирования был изменен с Delphi на C#.

Программный комплекс был разработан согласно модульной структуре с использованием классов. Модульная структура необходима для внедрения в будущем модулей, например, для создания фильтров, воспроизведения аудиозаписи и т.д.

Графический интерфейс программы состоит из двух окон: для выбора файла для обработки и для визуального представления построенных графиков. Архитектура комплекса представлена на рис. 1. Для

работы с графиками была добавлена функция ручного выделения участков графика.

Программный комплекс делится на три основных модуля: фильтрация, сегментация и построение графиков. В фильтрации данные из аудиофайла считываются в соответствующие переменные. Далее речевой сигнал сворачивается с двумя фильтрами и раскладывается на  $\sin$  и  $\cos$  составляющую. Следующим шагом рассчитываются интенсивности для 128 каналов для каждого отчета времени по заданной математической формуле. Все расчётные формулы представлены в [16]. В данном модуле после вычисления интенсивности вычисляется мгновенная частота и происходит одновременная маскировка с соответствующим фильтром.

Слуховая система человека обладает эффектом одновременной маскировки частот. Этот эффект возникает, когда рядом расположенные нейроны воспринимают две или более компоненты, частоты которых находятся недалеко друг от друга [17]. Процесс одновременной маскировки имитирует данный эффект.

После окончания работы модуля фильтрации обработанный сигнал переходит в модуль сегментации. Первый шаг в данном модуле заключается в вычислении канала ЧОТ. Далее производится сама сегментация сигнала, т.е. выделения вокализованных и невокализованных участков сигнала. Алгоритм сегментации представлен в [17].

После работы двух предыдущих модулей все полученные параметры поступают на вход процесса построения графиков. Графики представляют собой зависимость параметров от временных отчетов, например, мгновенной частоты от временных отчетов. На полученных графиках вручную можно разметить области и дать им название. Это сделано для ручного выделения фонем на графике для определения параметрического описания каждой фонемы.

#### **Описание структуры программного комплекса**

Структура комплекса представляет собой модули, которые состоят из классов, на рис. 2 представлена диаграмма классов в нотации UML. Модуль фильтрации включает в себя классы: `Header_Fltr`, `Fltr_flk`, `Fltrflt`, `File_Open`, `Header_Segm` и `Filter_Segm` – класс для сегментации сигнала, классы `Graph`, `Main_Window` и `Menu_Window` являются модулями для построения и отображения графиков.

Назначение классов программного комплекса:

– `Header_Fltr` – абстрактный класс, который содержит в себе структуру фильтров в виде атрибутов класса и также абстрактную функцию для считывания готовых фильтров.

– `Fltr_flk` и `Fltrflt` – классы для считывания  $\sin$  и  $\cos$  фильтров и фильтра для маскировки соответственно. Данные классы являются наследниками класса `Header_Fltr`.

– `File_Open` – класс для считывания данных из аудиофайла формата wav. При этом для работы с комплексом используется wav-файл со следующими параметрами: дискретизация – 12 кГц, качество звука – 16 бит, моносигнал.

- Header\_Segm – класс для работы с вычисляемыми параметрами.
- Filter – в данном классе происходит фильтрация речевого сигнала.
- Segm – в данном классе происходит сегментация речевого сигнала на вокализованные и невокализованные участки.
- Graph – класс для вычисления координат для графиков.

- Main\_Window и Menu\_Window – окна для построения графиков и меню соответственно.

В настоящее время используются готовые фильтры, файлы с фильтрами были вшиты в программу для удобства использования. В дальнейшем в комплекс будет добавлен модуль по созданию фильтров, чтобы улучшить процесс фильтрации речевого сигнала. При этом структура фильтра останется той же.

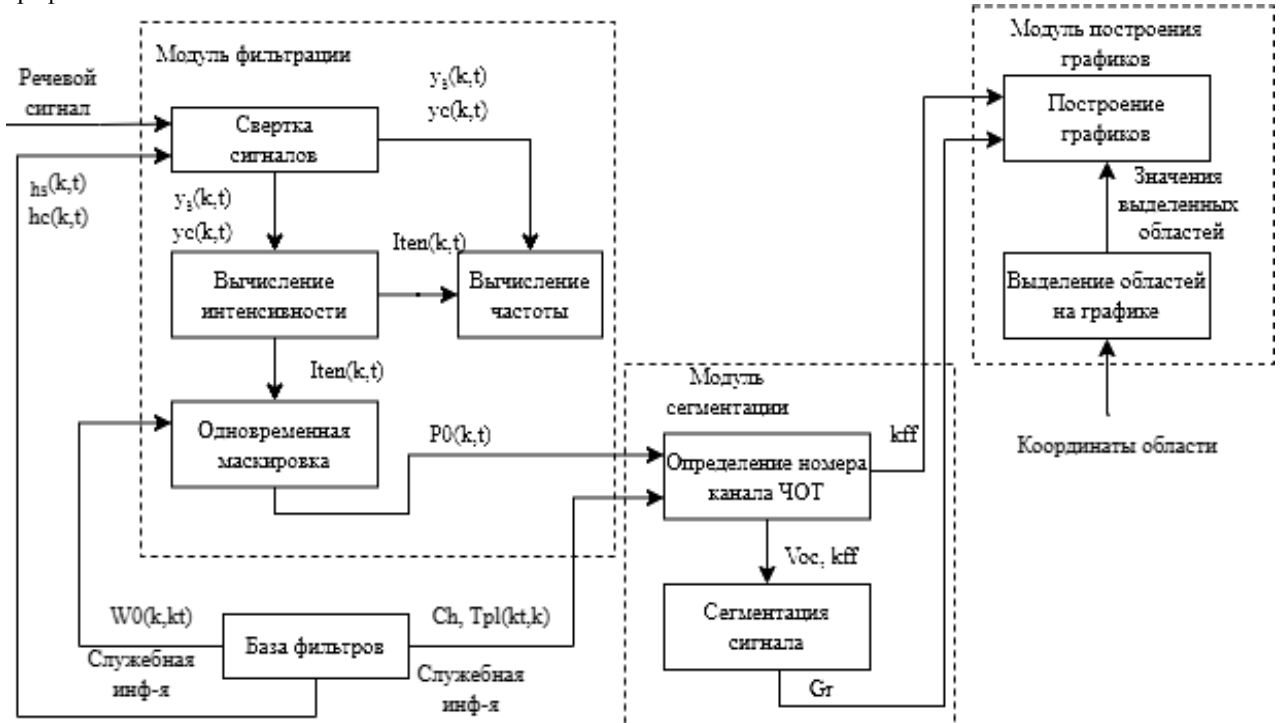


Рис. 1. Архитектура комплекса для исследования речевого сигнала

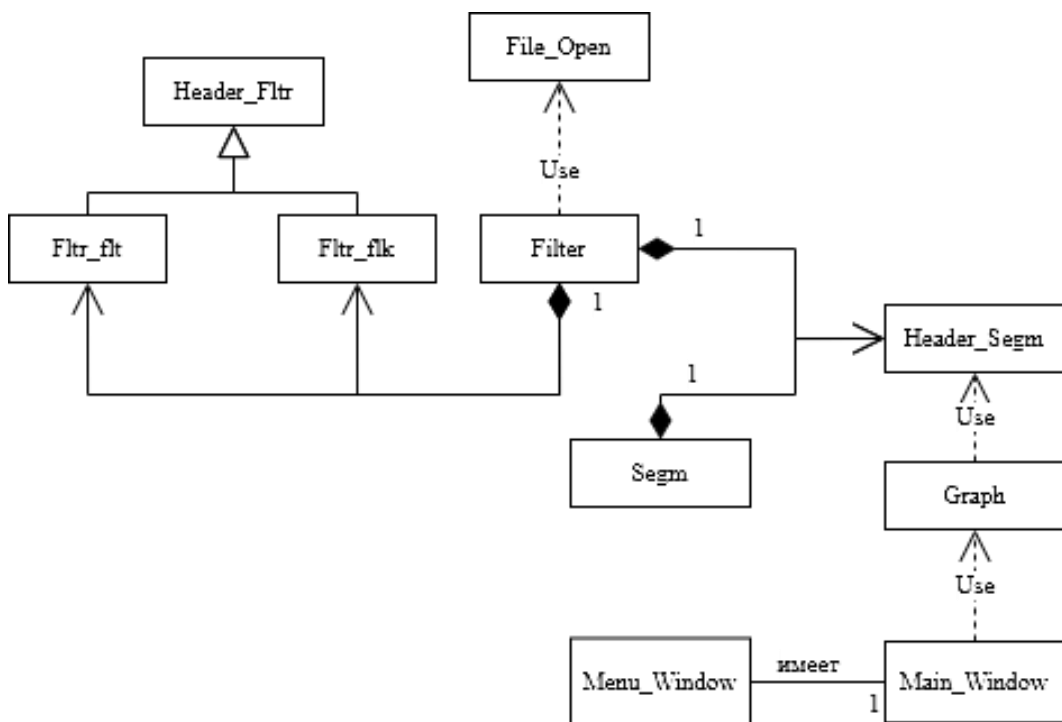


Рис. 2. Структура комплекса для исследования речевого сигнала

```

Структура фильтров:
public string sFlt; // "FLT ", "MFLT"
public short Channels; // количество каналов
public short SRate; // частота дискретизации
public short Fmin; // минимальная частота филь-
трации
public short Fmax; // максимальная частота филь-
трации
public short Qn; // добротность на нижней частоте
public short Qv; // добротность на верхней ча-
стоте
public short Qid; // тип задания добротности: 1 –
exp, 2 – линейно, 3 – по таблице
public char SinCos; // 'S'-синусная составляющая,
'C'-косинусная
public double a; // коэффициент точности для
нахождения длины окна
public short[] i; // количество выборок по оси вре-
мени / 2
public double[][] fltr; //массив для считывания
    
```

```

public string name { get; set; } // имя файла с филь-
тром
    
```

Представленная структура комплекса позволяет оптимально внедрять новые модули и модернизировать уже имеющиеся.

### Графический интерфейс программного комплекса

Для разработки дружелюбного интерфейса и удобного взаимодействия с графиками был использован фреймворк Avalonia. В отличие от классического приложения Windows.Forms, Avalonia позволяет разрабатывать кроссплатформенные приложения в том числе и для операционных систем с ядром Linux. Также у Avalonia функциональных возможностей больше для работы с графиками, имеются множество готовых библиотек. Для построения графиков была использована библиотека ScottPlot. На рис. 3 представлен графический интерфейс для построения графиков.

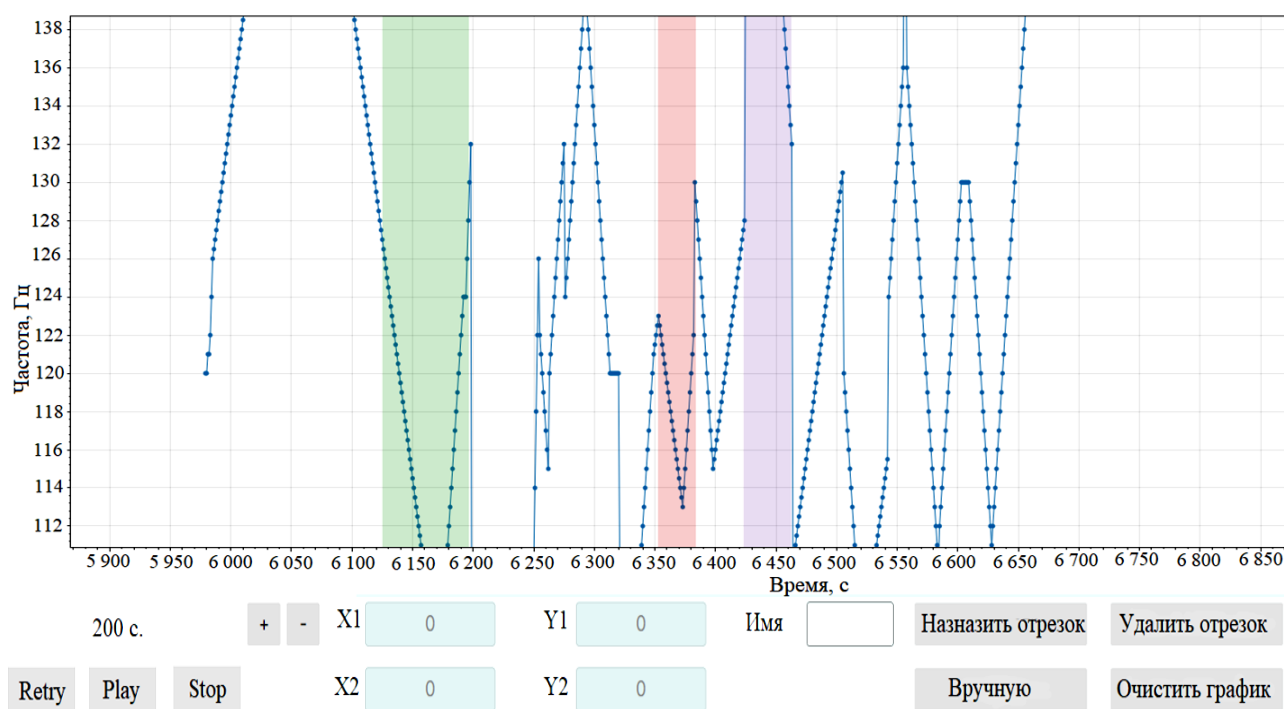


Рис. 3. Графический интерфейс для построения графиков: зависимость  $x$  от  $y$

На построенных графиках можно выделить различными цветами необходимые участки графика и назначить им имена (метки), например, параметрическое значение той или иной фонемы. При этом значения координат для выделяемых участков можно ввести вручную. Для удобства просмотра значений при наведении на какую-либо точку на графике выводится ее значение координат.

Назначенные отрезки можно удалить с графика, также есть возможность очистить график от всех меток. В данной версии программы аудиофайл проигрывается без звуковой дорожки, но есть возможность поставить на паузу или же начать воспроизведение заново. График масштабируется либо специальными кнопками, либо с помощью компьютерной мыши.

### Заключение

В работе описан программный комплекс для изучения параметров речевого сигнала, разработанный на основе работы [16]. Для модификации предыдущей версии комплекса были поставлены требования, с учетом которых разрабатывался данный комплекс. Подробно описаны архитектура, структура и графический интерфейс программы. Преимущества модернизированного программного комплекса от предыдущей версии – это модульная архитектура и графический интерфейс, который позволяет работать с графиками, в том числе вручную выделять области графика, что повышает эффективность взаимодействия пользователя с системой.

Дальнейшая модернизация программного комплекса предполагает внедрение новых модулей, описанных в данной работе, и доработку реализованных функций при выявлении ошибок в работе программы.

Работа выполнена при финансовой поддержке Министерства науки и высшего образования РФ в рамках базовой части государственного задания ТУ-СУРА на 2023–2025 гг. (проект № FEWM-2023-0015).

### Литература

1. Анализ рынка речевых технологий и распознавания речи в мире [Электронный ресурс]: сайт Fortune Business Insights. – URL: <https://www.fortunebusinessinsights.com/industry-reports/speech-and-voice-recognition-market-101382>, свободный (дата обращения: 10.11.2024).

2. Bhatt S. Continuous Speech Recognition Technologies – A Review / S. Bhatt, A. Jain, A. Dev // *Recent Developments in Acoustics*. – 2021. – P. 85–94.

3. Катаев М.Ю. Методика распознавания речевых команд в школьных информационных системах // *Речевые технологии*. – 2024. – № 1. – С. 18–34.

4. Truslow E. KlattWare tools for acoustic analysis of speech signals / E. Truslow, H.M. Hanson // *The Journal of the Acoustical Society of America*. – 2010. – Vol. 128, No. 4. – P. 2290.

5. Martin P. WinPitchPro. A Tool for Text to Speech Alignment and Prosodic Analysis // *Speech Prosody*. – 2004. – P. 1–4.

6. Haubold A. Alignment of speech to highly imperfect text transcriptions / A. Haubold, J.R. Kender // *2007 IEEE International Conference on Multimedia and Expo*. – 2007. – P. 1–4. DOI: 10.1109/ICME.2007.4284627.

7. Mohammed A.A. Fundamental Frequency and Jitter Percent in MDVP and PRAAT / A.A. Mohammed, A. Nagy // *J. Voice*. – 2023. – Vol. 37, No. 4. – P. 496–503.

8. Oğuz H. Comparison of results in two acoustic analysis programs: Praat and MDVP / H. Oğuz, M.A. Kiliç, M.A. Şafak // *Turkish Journal of Medical Sciences*. – 2011. – Vol. 41, No. 5. – P. 835–841.

9. Voice disorders in severe obstructive sleep apnea patients and comparison of two acoustic analysis software programs: MDVP and Praat / M. Wei, J. Du, X. Wang, H. Lu, W. Wang, P. Lin // *Sleep Breath*. – 2021. – Vol. 25, No. 1. – P. 433–439.

10. Perturbation measures of voice: a comparative study between Multi-Dimensional Voice Program and Praat / Y. Maryn, P. Corthals, M. De Bodt, P. Van Cauwenberge, D. Deliyiski // *Folia Phoniatr Logop*. – 2009. – Vol. 61, No. 4. – P. 217–226.

11. Ofer A. A clinical comparison between two acoustic analysis softwares: MDVP and Praat // *Biomedical Signal Processing and Control*. – 2009. – Vol. 4, No. 3. – P. 202–205.

12. Multidimensional Voice Program (MDVP) and amplitude variation parameters in euphonic adult subjects. Normative study / M. Nicastrì, G. Chiarella, L.V. Gallo, M. Catalano, E. Cassandro // *Acta otorhinolaryngol Ital*. – 2004. – Vol. 24. – P. 337–341.

13. A Comparison of Healthy and Disordered Voices Using Multi-Dimensional Voice Program, Praat, and TF32 / L.-C. Keung, K. Richardson, D. S. Matheron, V. Martel-Sauvageau // *Journal of Voice*. – 2024. – Vol. 38, No. 4. – P. 23–38.

14. Мырадов М.Т. Современные технологии распознавания речи / М.Т. Мырадов, С.Н. Назарова // *Наука и мирознание*. – 2024. – № 1 (23). – С. 301–305.

15. Якимук А.Ю. Программный комплекс для автоматизации моделирования сегментации речевых сигналов и вокальных исполнений / А.Ю. Якимук, А.А. Конев, А.О. Осипов // *Вестник Иркут. гос. техн. ун-та*. – 2017. – Т. 21, № 10 (129). – С. 53–64.

16. Конев А.А. Модель и алгоритмы анализа и сегментации речевого сигнала: автореф. дис. ... канд. техн. наук. – Томск, 2007. – 20 с.

17. Конев А.А. Сегментация речевых сигналов на вокализованные и невокализованные участки на основе одновременной маскировки / А.А. Конев, Р. В. Мещеряков, Е.Ю. Костюченко // *Автометрия*. – 2018. – Т. 54, № 4. – С. 51–57.

### Репьюк Наталья Сергеевна

Аспирант каф. комплексной информационной безопасности электронно-вычислительных систем (КИБЭВС) Томского государственного университета систем управления и радиоэлектроники (ТУСУР) Ленина пр-т, 40, г. Томск, Россия, 634050  
Тел.: +7 (382-2) 70-15-29  
Эл. почта: rns@fb.tusur.ru

### Конев Антон Александрович

Канд. техн. наук, доцент, доцент каф. КИБЭВС ТУСУР  
Ленина пр-т, 40, г. Томск, Россия, 634050  
ORCID: 0000-0002-3222-9956  
Тел.: +7 (382-2) 70-15-29  
Эл. почта: kaa@fb.tusur.ru

Поступила в редакцию: 23.12.2024.

Принята к публикации: 21.04.2025.

Repyuk N.S., Konev A.A.

### Software package for the study of speech signals

This paper presents a software package designed to study the characteristics of speech signal parameters. This complex is planned to be used to develop a parametric description of sounds and (or) groups of sounds. The presented complex was developed on the basis of algorithms developed by A.A. Konev, taking into account the requirements for modifying the previous version of the complex. The structure, architecture and main blocks of this complex are described.

**Keywords:** software package, speech signal, filtering, segmentation.

**DOI:** 10.21293/1818-0442-2025-28-1-93-99

### References

1. *Analiz rynka rechevyh tehnologij i raspoznavanija rechi v mire* [Analysis of speech recognition technologies in the world]. Available at: <https://www.fortunebusinessinsights.com/industry-reports/speech-and-voice-recognition-market-101382> (Accessed: November 10, 2024).

2. Bhatt S., Jain A., Dev A. Continuous Speech Recognition Technologies – A Review. *Recent Developments in Acoustics*, 2021, pp. 85–94.

3. Kataev M.Ju. [Methods of speech command recognition in school information systems]. *Speech Technology*, 2024, No. 1, pp. 18–34 (in Russ.).

4. Truslow E., Hanson H.M. KlattWare tools for acoustic analysis of speech signals. *The Journal of the Acoustical Society of America*, 2010, vol. 128, No. 4, pp. 2290.
5. Martin P. WinPitchPro. A Tool for Text to Speech Alignment and Prosodic Analysis. *Speech Prosody*, 2004, pp. 1–4.
6. Haubold A., Kender J.R. Alignment of speech to highly imperfect text transcriptions. *2007 IEEE International Conference on Multimedia and Expo*, 2007, pp. 1–4. DOI: 10.1109/ICME.2007.4284627.
7. Mohammed A.A., Nagy A. Fundamental Frequency and Jitter Percent in MDVP and PRAAT. *Journal of Voice*, 2023, vol. 37, No. 4, pp. 496–503.
8. Oğuz H., Kiliç M.A., Şafak M.A. Comparison of results in two acoustic analysis pro-grams: Praat and MDVP. *Turkish Journal of Medical Sciences*, 2011, vol. 41, No. 5, pp. 835–841.
9. Wei M., Du J., Wang X., Lu H., Wang W., Lin P. Voice disorders in severe obstructive sleep apnea patients and comparison of two acoustic analysis software pro-grams: MDVP and Praat. *Sleep Breath*, 2021, vol. 25, No. 1, pp. 433–439.
10. Maryn Y., Corthals P., De Bodt M., Van Cauwenberge P., Deliyski D. Perturbation measures of voice: a comparative study between Multi-Dimensional Voice Program and Praat. *Folia Phoniatica et Logopaedica*, 2009, vol. 61, No. 4, pp. 217–226.
11. Ofer A. A clinical comparison between two acoustic analysis softwares: MDVP and Praat. *Biomedical Signal Processing and Control*, 2009, vol. 4, No. 3, pp. 202–205.
12. Nicastrì M., Chiarella G., Gallo L.V., Catalano M., Cassandro E. Multidimensional Voice Program (MDVP) and amplitude variation parameters in euphonic adult subjects. Normative study. *ACTA Otorhinolaryngologica Italica*, 2004, vol. 24, pp. 337–341.
13. Keung L.-C., Richardson K., Matheron D.S., Martel-Sauvageau V. A Comparison of Healthy and Disordered Voices Using Multi-Dimensional Voice Program, Praat, and TF32. *Journal of Voice*, 2024, vol. 38, No. 4, pp. 23–38.
14. Muradov M.T., Nazarova S.N. [Modern speech recognition technologies]. *Science and Worldview*, 2024, No. 1(23), pp. 301–305 (in Russ.).
15. Jakimuk A.Ju., Konev A.A., Osipov A.O. [Program complex for speech signal and vocal performance segmentation modeling automation]. *Proceedings of Irkutsk State Technical University*, 2017, vol. 21, No. 10(129), pp. 53–64 (in Russ.).
16. Konev A.A. *Model i algoritmy analiza i segmentacii rechevogo signala*. [Speech signal analysis and segmentation model and algorithms], *Dissertation for the Candidate of Science degree*, Tomsk, 2007. 20 p. (in Russ.).
17. Konev A.A., Meshherjakov R.V., Kostjuchenko E.Yu. [Segmentation of speech signals into vocalized and non-vocalized areas based on simultaneous masking]. *Avtometriya*, 2018, vol. 54, no 4, pp. 51–57 (in Russ.).

**Natalia S. Repyuk**

Postgraduate student, Department of Complex Information Security of Computer Systems, Tomsk State University of Control Systems and Radioelectronics (TUSUR)  
40, Lenin pr., Tomsk, Russia, 634050  
Phone: +7 (382-2) 70-15-29  
Email: rns@fb.tusur.ru

**Anton A. Konev**

Candidate of Sciences in Engineering, Assistant Professor, Department of Complex Information Security of Computer Systems, TUSUR  
40, Lenin pr., Tomsk, Russia, 634050  
ORCID: 0000-0002-3222-9956  
Phone: +7 (382-2) 70-15-29  
Email: kaa@fb.tusur.ru

Received: 23.12.2024.

Accepted: 21.04.2025.