

## Список использованных источников

1. Методы и средства научных исследований: учеб. пособие / Ю.Н. Колмогоров [и др.]. — Екатеринбург: Изд-во Урал. ун-та, 2017. — 152 с.

УДК 004.934:004.056

**И.А. Третьяков, Е.Н. Кожекина**  
Донецкий национальный университет  
Донецк

### РЕАЛИЗАЦИЯ И ТЕСТИРОВАНИЕ АВТОМАТИЗИРОВАННОЙ СИСТЕМЫ ТЕКСТОНЕЗАВИСИМОЙ ИДЕНТИФИКАЦИИ ДИКТОРА

*Аннотация.* Реализована автоматизированная система текстонезависимой идентификации говорящего, применяющая методы мел-частотных кепстральных коэффициентов и коэффициентов линейного предсказания. Оценена точность применяемых методов.

**I.A. Tretiakov, E.N. Kozhekina**  
Donetsk National University  
Donetsk

### IMPLEMENTATION AND TESTING OF AN AUTOMATED TEXT-INDEPENDENT SPEAKER IDENTIFICATION SYSTEM

*Abstract.* An automated system of text-independent speaker identification is implemented, using the methods of low-frequency kepsral coefficients and linear prediction coefficients. The accuracy of the applied methods is estimated.

Голосовая идентификация является частью отдельного научного направления – теории речеобразования [1-3]. Перспективным представляется ее новое применение в автоматизированных системах, основанных на телекоммуникационных каналах связи [4-5]. В качестве примера, в мобильной связи с помощью голоса можно осуществлять управление услугами, причем внедрение голосовой идентификации способствует защите от мошенничества.

Преимуществами голосовой биометрии является низкая стоимость и простота внедрения. К недостаткам относится невысокая надежность по сравнению с остальными методами. Для повышения

надежности в некоторых системах используется текстозависимая идентификация, например, когда пользователю предлагается произнести случайно сгенерированную парольную фразу. Текстонезависимая идентификация подразумевает использование только индивидуальных признаков.

Процедура текстонезависимой идентификации в данной работе состоит из четырех этапов, которые более подробно описаны в [5]:

- этап получения сигнала;
- этап предварительной обработки;
- этап извлечения признаков;
- этап классификации признаков.

Текстонезависимая идентификация диктора была реализована на языке программирования Python 3.6. Выбор обоснован простотой синтаксиса и большим набором библиотек. В данной работе при выборе библиотек учитывалась скорость работы, использование вычислительных мощностей, а также простота использования. В программу вошли следующие библиотеки: Numpy, Scipy, python\_speech\_features, Wav, pyAudio, sklearn. Извлечение признаков осуществляется при помощи методов MFCC (мел-частотные кепстральные коэффициенты) и LPC (коэффициенты линейного предсказания). В качестве классификатора была использована модель гауссовых смесей.

Для проведения тестирования разрабатываемой системы в работе был использован набор голосовых данных «Common Voice Corpus» (многоязычная коллекция звуковых файлов, предназначенная для исследований и разработок в области речевых технологий) на русском языке. Из данной коллекции были случайным образом выбраны 5 дикторов (голосов). Речь каждого диктора была разбита на 15 фрагментов, длительностью от 2 до 10 секунд. Особенности дикторов: дикторы 1, 2 и 3 имеют голос человека подросткового возраста, причем диктор 1 женского пола (диктор 2 и 3 мужского). Дикторы 4 и 5 имеют голоса соответственно взрослой женщины и взрослого мужчины.

Используемые фрагменты речи имеют формат mp3. Разрабатываемая система использует библиотеку scipy.io.wavfile, поэтому требует формат wav. Для повышения точности тестирования, все аудио-фрагменты подвергались обработке в программе Audacity, с помощью которой из звуков удалялись участки тишины и паузы, возникающие при произнесении предложений.

Тестирование осуществляется идентификацией неизвестного диктора путем сравнения максимального значения правдоподобия для каждого диктора. Эффективность идентификации определяется по

общему показателю точности в условиях белого шума и речеподобной помехи. При проведении тестирования уровень громкости этих шумов будет постепенно увеличиваться, что позволит выявить сильные и слабые стороны рассматриваемых методов. Изменение громкости шумов производится на этапе смешивания образца звука с шумом при помощи библиотеки pyAudio.

Результаты тестирования метода MFCC (мел-частотных кепстральных коэффициентов) при влиянии белого шума показаны на рис. 1, при влиянии речеподобной помехи показаны на рис. 2.

Точность идентификации для speaker 1 : 100.0%	
Точность идентификации для speaker 2 : 100.0%	
Точность идентификации для speaker 3 : 100.0%	
Точность идентификации для speaker 4 : 86.66666666666667%	10
Точность идентификации для speaker 5 : 100.0%	дБ
Общая точность идентификации: 97.33333333333334%	
Время выполнения: 6735 мс.	
Точность идентификации для speaker 1 : 100.0%	
Точность идентификации для speaker 2 : 93.33333333333333%	
Точность идентификации для speaker 3 : 100.0%	
Точность идентификации для speaker 4 : 33.33333333333336%	15
Точность идентификации для speaker 5 : 100.0%	дБ
Общая точность идентификации: 85.33333333333333%	
Время выполнения: 6173 мс.	
Точность идентификации для speaker 1 : 100.0%	
Точность идентификации для speaker 2 : 6.666666666666665%	
Точность идентификации для speaker 3 : 0.0%	
Точность идентификации для speaker 4 : 0.0%	25
Точность идентификации для speaker 5 : 100.0%	дБ
Общая точность идентификации: 41.33333333333336%	
Время выполнения: 6185 мс.	

**Рис. 1- Влияние белого шума на идентификацию**

Точность идентификации для speaker 1 : 100.0%	
Точность идентификации для speaker 2 : 100.0%	
Точность идентификации для speaker 3 : 100.0%	
Точность идентификации для speaker 4 : 80.0%	10
Точность идентификации для speaker 5 : 100.0%	дБ
Общая точность идентификации: 96.0%	
Время выполнения: 6116 мс.	
Точность идентификации для speaker 1 : 100.0%	
Точность идентификации для speaker 2 : 100.0%	
Точность идентификации для speaker 3 : 73.33333333333334%	
Точность идентификации для speaker 4 : 19.999999999999996%	15
Точность идентификации для speaker 5 : 100.0%	дБ
Общая точность идентификации: 78.66666666666666%	
Время выполнения: 6048 мс.	

Точность идентификации для speaker 1 : 100.0%	
Точность идентификации для speaker 2 : 100.0%	
Точность идентификации для speaker 3 : 0.0%	
Точность идентификации для speaker 4 : 0.0%	25
Точность идентификации для speaker 5 : 80.0%	дБ
Общая точность идентификации: 56.00000000000001%	
Время выполнения: 6694 мс.	

**Рис. 2 - Влияние речеподобной помехи на идентификацию**

Результаты тестирования метода LPC (коэффициентов линейного предсказания) при влиянии белого шума показаны на рис. 3, при влиянии речеподобной помехи показаны на рис. 4.

Точность идентификации для speaker 1 : 100.0%	
Точность идентификации для speaker 2 : 100.0%	
Точность идентификации для speaker 3 : 100.0%	
Точность идентификации для speaker 4 : 100.0%	15
Точность идентификации для speaker 5 : 100.0%	дБ
Общая точность идентификации: 100.0%	
Точность идентификации для speaker 1 : 100.0%	
Точность идентификации для speaker 2 : 100.0%	
Точность идентификации для speaker 3 : 66.66666666666667%	25
Точность идентификации для speaker 4 : 100.0%	
Точность идентификации для speaker 5 : 93.33333333333333%	дБ
Общая точность идентификации: 92.0%	

**Рис. 3 - Влияние белого шума на идентификацию**

Точность идентификации для speaker 1 : 100.0%	
Точность идентификации для speaker 2 : 100.0%	
Точность идентификации для speaker 3 : 100.0%	
Точность идентификации для speaker 4 : 100.0%	15
Точность идентификации для speaker 5 : 93.33333333333333%	дБ
Общая точность идентификации: 98.66666666666667%	
Точность идентификации для speaker 1 : 93.33333333333333%	
Точность идентификации для speaker 2 : 40.0%	
Точность идентификации для speaker 3 : 0.0%	
Точность идентификации для speaker 4 : 73.33333333333334%	25
Точность идентификации для speaker 5 : 6.666666666666665%	дБ
Общая точность идентификации: 42.666666666666664%	

**Рис. 4 - Влияние речеподобной помехи на идентификацию**

Результаты эксперимента показали, что оба метода подвержены влиянию шумов. Несмотря на то, что метод LPC показал невысокую точность идентификации, во всех остальных тестах общая точность

идентификации составляла более чем 90%. Метод LPC подвержен влиянию только речеподобной помехи, в то время как метод MFCC подвержен влиянию всех видов помех. Однако влияние от белого шума и речеподобной помехи оказались примерно на одном уровне.

Таким образом, анализ результатов тестовых экспериментов показал, что наиболее подходящим методом извлечения признаков для дальнейшей идентификации диктора в рамках данной работы является метод коэффициентов линейного предсказания (LPC).

### **Список использованных источников**

1. Фант, Г. Акустическая теория речеобразования / Г. Фант. – М.: Наука, 1964. – 283 с.
2. Меньшаков, П. А. Голосовая идентификация в системах контроля доступа / П. А. Меньшаков, И. А. Мурашко // Вестник ХГУ им. Н.Ф. Катанова. – 2017. – №20. – С. 34-37.
3. Наздрюхин, А. С. Фильтрация сибилантов для улучшения качества речевого сигнала / А. С. Наздрюхин, Г. А. Левченко // Вопросы развития современной науки и техники. – 2021. – № 5. – С. 165-170.
4. Третьяков, И. А. Исследование параметров рекуррентной нейронной сети для распознавания человека по голосу в системах безопасности / И. А. Третьяков, Е. Н. Кожекина, А. Е. Мышкин // Вестник Донецкого национального университета. Серия Г: Технические науки. – 2021. – № 4. – С. 24-36.
5. Третьяков, И. А. Текстонезависимая идентификация речи в условиях помех / И. А. Третьяков, Е. Н. Кожекина, В. И. Сыровацкий // Вестник Донецкого национального университета. Серия Г: Технические науки. – 2022. – № 2. – С. 64-77.

УДК 330.36

**В. А. Турко**

Центр системного анализа и стратегических исследований  
Национальной академии наук Беларуси  
Минск, Беларусь

### **ГОСПЛАН ПРОВЕРЕННЫЙ ИНСТРУМЕНТ ЭКОНОМИЧЕСКОГО СУВЕРЕНИТЕТА**

*Аннотация.* Обоснована необходимость наличия качественного инструментария для сбалансированного развития многоотраслевого комплекса