

ПРОГРАММНО-АППАРАТНАЯ РЕАЛИЗАЦИЯ СИСТЕМЫ ОБНАРУЖЕНИЯ И КЛАССИФИКАЦИИ ЭМОЦИОНАЛЬНЫХ СОСТОЯНИЙ ЧЕЛОВЕКА ПО РЕЧЕВЫМ СИГНАЛАМ В ЭКСТРЕМАЛЬНЫХ УСЛОВИЯХ

А. К. Алимуратов

Пензенский государственный университет, Пенза, Россия
alansapfir@yandex.ru

Аннотация. *Актуальность и цели.* Повышение достоверности обнаружения и классификации критических эмоциональных состояний в отраслях человеческой деятельности, связанных с риском возникновения техногенных и биогенных аварий становится все более актуальной задачей. Целью исследования является повышение эффективности системы обнаружения и классификации эмоциональных состояний человека по речевым сигналам. Объектом исследования являются структура и программно-аппаратная реализация системы обнаружения и классификации эмоциональных состояний. Предметом исследования являются средства и технические решения программной и аппаратной реализации. *Материалы и методы.* В рамках исследования использовались методы клиент-серверной программно-аппаратной реализации системы на основе оригинальных программных модулей обработки речевых сигналов и классификации эмоциональных расстройств, а также нескольких аппаратных реализаций серверного устройства. *Результаты.* В статье представлена программно-аппаратная реализация системы обнаружения и классификации критических эмоциональных состояний человека по речевым сигналам в виде клиент-серверного решения в экстремальных условиях. Оригинальность предложенной концепции заключается в создании кластеров серверных устройств, связанных между собой высокоскоростным каналом. Данный подход к организации клиент-серверного решения системы позволяет повысить производительность, увеличить скорость обработки данных, обеспечить стабильность работы и своевременную выдачу результатов в режиме реального времени. *Выводы.* В соответствии с полученными результатами исследования выявлено, что достоверность разработанной системы для классификации эмоциональных состояний составляет 94,7–95,6 %, для классификации эмоций 93,1–95,6 %. В сравнении с аналогичными продуктами на рынке речевых технологий достоверность разработанной системы в среднем больше на 3,15 и 2,35 % для классификации эмоциональных состояний и эмоций соответственно. В перспективе планируется провести дополнительные исследования быстродействия программно-аппаратной реализации системы обнаружения и классификации критических эмоциональных состояний человека по речевым сигналам в экстремальных условиях.

Ключевые слова: обработка речевых сигналов, эмоциональная речь, обнаружение и классификация эмоциональных состояний человека, клиент-серверное решение

Для цитирования: Алимуратов А. К. Программно-аппаратная реализация системы обнаружения и классификации эмоциональных состояний человека по речевым сигналам в экстремальных условиях // Измерение. Мониторинг. Управление. Контроль. 2023. № 2. С. 116–128. doi: 10.21685/2307-5538-2023-2-14

SOFTWARE AND HARDWARE IMPLEMENTATION OF A SYSTEM TO DETECT AND CLASSIFY HUMAN EMOTIONAL STATES FROM SPEECH SIGNALS

A.K. Alimuradov

Penza State University, Penza, Russia
alansapfir@yandex.ru

Abstract. *Background.* To improve the accuracy of detection and classification of emotional states in the areas of human activity associated with the risk of man-made and biogenic accidents has become an urgent task. The aim of the study is to improve the efficiency of a system to detect and classify human emotional states from speech signals. The object of the research is the structure and software and hardware implementation of the system for detecting and classifying

emotional states. The subject of the study is the means and techniques for software and hardware implementation. *Materials and methods.* The client-server methods for hardware and software implementation of the system based on the original software modules for speech signal processing and classifying emotional disorders, as well as several hardware implementations of the server device have been used. *Results.* The article presents hardware and software implementation of a system to detect and classify human emotional states from speech signals based on the client-server architecture. The novelty of the proposed concept is the creation of server clusters interconnected by a high-speed channel. This approach to the organization of the client-server architecture of the system allows increasing productivity and the speed of data processing, ensuring stability and timely delivery of results in real time. *Results and conclusions.* The research results have evidenced the accuracy of the developed system for the classification of emotional states being 94.7–95.6 %, and that for the classification of emotions being 93.1–95.6 %. The accuracy of the developed system is on average 3.15 % and 2.35 % higher for the classification of emotional states and emotions, respectively, as compared with related products on the speech technology market. In the future, it is planned to conduct further research on the performance of hardware and software implementation of the system to detect and classify human emotional states from speech signals.

Keywords: speech signal processing, emotional speech, detection and classification of human emotional states, client-server architecture

For citation: Alimuradov A.K. Software and hardware implementation of a system to detect and classify human emotional states from speech signals. *Izmerenie. Monitoring. Upravlenie. Kontrol'* = *Measuring. Monitoring. Management. Control.* 2023;(2):116–128. (In Russ.). doi: 10.21685/2307-5538-2023-2-14

Введение

В настоящее время обнаружение и классификация критических эмоциональных состояний в отраслях человеческой деятельности, связанных с риском возникновения техногенных и биогенных аварий, террористических угроз с серьезными последствиями, становится все более актуальной проблемой. Роль достоверности обнаружения «человеческого фактора» значительно возрастает, если трудовая деятельность сопровождается сложными когнитивными процессами, дефицитом времени и высокой эмоциональной нагрузкой.

На сегодняшний день в области речевой аналитики имеются технологии и продукты, предназначенные для обнаружения и классификации эмоциональных состояний человека по речевым сигналам в экстремальных условиях. В табл. 1 представлена информация о системах классификации и распознавания эмоциональных состояний/эмоций, получивших широкую практическую популярность. В рамках анализа информации в табл. 1 наибольший интерес представляют методы и способы обработки эмоциональных речевых сигналов, а также программно-аппаратная реализация системы классификации отрицательных, нейтральных и положительных эмоциональных состояний человека. Однако в силу коммерческой тайны производители не предоставляют данную информацию. Информация в табл. 1 не претендует на полноту и демонстрирует многообразие, достоинства и недостатки аналогичных продуктов на рынке речевых технологий. Подробное описание перечисленных выше продуктов можно найти на сайтах производителей [1–6].

В соответствии с результатами сравнительного анализа сделаны следующие выводы:

– существует необходимость совершенствования и разработки новых методов и средств обработки и анализа естественной эмоциональной речи, обеспечивающих повышение достоверности фильтрации, сегментации и определения информативных параметров, релевантных эмоциональным состояниям;

– существует необходимость разработки оригинальной концепции программно-аппаратной реализации системы, позволяющей повысить достоверность и снизить уровень ошибок обнаружения и классификации критических эмоциональных состояний человека по речевым сигналам.

В данной статье представлена программно-аппаратная реализация системы обнаружения и классификации критических эмоциональных состояний человека по речевым сигналам в виде клиент-серверного решения. Оригинальность предложенной концепции заключается в использовании нескольких аппаратных реализаций серверного устройства – создание кластеров, связанных между собой высокоскоростным каналом. Данный подход к организации клиент-серверного решения системы позволяет повысить производительность, увеличить скорость обработки данных, обеспечить стабильность работы и своевременную выдачу результатов в режиме реального времени в экстремальных условиях.

Таблица 1

Системы обнаружения и классификации эмоциональных состояний человека по речевым сигналам, представленные на рынке речевых технологий

Наименование. Тип доступа	Язык(и)	Виды эмоций / характер выражения эмоций	Достоверность обнаружения и классификации	Функционал/назначение	Производитель / Стоймость	Условия регистрации речевых данных
1	2	3	4	5	6	7
1. Система речевой аналитики «SVQA» [1]. Тип доступа: закрытый, платный	Английский, русский, белорусский	Положительные, отрицательные и нейтральные эмоции / вынужденные эмоции	Не менее 90 %	Оценка качества речи, оценка эмоциональной составляющей речи / Безопасность, бизнес (банковская сфера, контакт-центры)	Частотное предприятие «Речевые системы» https://svqa.ru Беларусь, г. Минск/ По запросу	Телефонный канал, записанные фонограммы
2. Подсистема автоматической оценки качества «Smart Logger» многоканальной системы регистрации телефонных вызовов и речевых сообщений «НЕЗАБУДКА П» [2]. Тип доступа: закрытый, платный	Английский, русский,	Радость, удивление, грусть, страх, гнев, отвращение / вынужденные и естественные эмоции	Ошибка 1-го рода – не более 10 %, ошибка 2-го рода – не более 5 %	Распознавание речи, идентификация личности по речи, поиск ключевых слов, оценка эмоциональной составляющей речи / Безопасность, бизнес (банковская сфера, контакт-центры)	Группа компаний «Центр речевых технологий» (ЦРТ) (компания ЦРТ, ЦРТ-инновации и SpeechPro) https://www.speechpro.ru Россия, г. Санкт-Петербург / По запросу	Телефонный канал, записанные фонограммы
3. Профессиональная система диагностики голоса и речи «ling WA VES» [3]. Тип доступа: закрытый, платный	Английский, немецкий, французский, испанский, итальянский	Радость, удивление, грусть, страх, гнев, отвращение / естественные эмоции	–	Речевая и языковая терапия, анализ эмоционального состояния, отоларингология / ЛОР-терапия, терапия для профессиональных певцов и ораторов/медицина	Компания WEVOSYS medical technology GmbH https://www.wevosys.com (Германия, г. Баунах) / По запросу	Специализированное микрофонное оборудование для регистрации речевых сигналов
4. Технология анализа голоса «Sense» [4]. Тип доступа: закрытый, платный	Английский, иврит	Радость, удивление, грусть, страх, гнев, отвращение / вынужденные и естественные эмоции	Не менее 92 %	Оценка качества речи, оценка эмоциональной составляющей речи / Безопасность, бизнес (банковская сфера, контакт-центры)	Компания «NemesYS CO» https://www.nemesysco.com (Израиль) / По запросу	Телефонный канал, записанные фонограммы

Окончание табл. 1

1	2	3	4	5	6	7
5. Технологии для анализа эмоций по голосу «Emprath» / «Web Emprath API» / «SDK Emprath API» / «Beluga Vox» / «Emo Value Generator» [5]. Тип доступа: закрытый, платный	Английский, японский	Положительные, отрицательные и нейтральные эмоции / вынужденные и естественные эмоции	–	Распознавание и генерация речи, анализ и оценка эмоциональной составляющей речи / Бизнес (банковская сфера, контакт-центры)	Компания «Emprath Inc.» https://www.webemprath.com (Япония, г. Токио) / По запросу	Телефонный канал, записанные фонограммы
6. Технологии аналитики речи «SmartSpeech» [6]. Тип доступа: закрытый, платный	Русский	Положительные, отрицательные и нейтральные эмоции / вынужденные и естественные эмоции	Не менее 90 %	Распознавание и генерация речи, анализ и оценка эмоциональной составляющей речи / Бизнес (банковская сфера, контакт-центры)	Компания ПАО «Сбербанк» https://sberdevices.ru Россия, г. Москва / По запросу	Телефонный канал, записанные фонограммы

Данная научная статья подготовлена в рамках проекта № МД-1066.2022.4 «Исследование скрытых паттернов речевых сигналов и разработка способов обнаружения и классификации естественно выраженных психоэмоциональных состояний человека», финансируемого Советом по грантам Президента РФ. Статья является продолжением ранее опубликованных научных работ [7–9].

Технические требования к программно-аппаратной реализации системы

Программная реализация системы представлена в виде мобильного приложения обнаружения и классификации эмоциональных отрицательных (отвращение, страх, гнев, печаль, удивление), нейтральных (нейтралитет) и положительных (удивление, радость) состояний человека по речевым сигналам в виде клиент-серверного решения.

Система должна состоять из модуля обработки речевых сигналов и анализа информативных параметров; модуля классификации, обученного на информативных параметрах, релевантных отрицательным, положительным и нейтральным эмоциональным состояниям.

Для полноценного функционирования мобильного приложения клиентского устройства системы требуются:

- процессор с тактовой частотой – не менее 1 ГГц;
- объем свободной оперативной памяти – не менее 512 Мб;
- объем свободной памяти на жестком диске – не менее 1 Гб.

Мобильное приложение системы должно обеспечивать обнаружение и классификацию критических эмоциональных состояний человека по речевым сигналам в режиме реального времени и функционировать в режиме доступа к сети «Интернет».

В целях обеспечения возможности реализации мобильного приложения в виде клиент-серверного решения и организации взаимодействия с несколькими устройствами (клиентами) одновременно требуется выполнение следующих требований:

- поддержка многопоточности у процессора серверного устройства;
- увеличенный объем оперативной памяти клиентского и серверного устройств.

Требование к многопоточности обосновывается тем, что для каждого взаимодействующего устройства необходимо выделять отдельное вычислительное ядро. Увеличенный объем оперативной памяти позволит хранить следующую информацию:

- данные с подключенных клиентских устройств к серверному устройству;
- данные для обеспечения работы системы в реальном времени.

Аппаратная реализация системы представлена в виде серверного устройства, включающего в себя: материнскую (системную) плату, процессор, оперативную память, графический ускоритель, блок питания, систему хранения данных и систему охлаждения.

Программно-аппаратная реализация системы

На рис. 1 представлена оригинальная структура клиент-серверного решения системы обнаружения и классификации критических эмоциональных состояний человека по речевым сигналам в экстремальных условиях, включающей в себя:

- подсистему обработки речевых сигналов;
- подсистему обнаружения и классификации эмоциональных состояний;
- подсистему управления взаимодействием процессами системы обнаружения и классификации эмоциональных состояний человека.

Клиентское устройство, осуществляющее регистрацию речевых сигналов, представляет собой персональное устройство Laptop/Desktop под управлением операционных систем Windows, Linux или macOS, в том числе мобильное устройство под управлением операционных систем iOS или Android.

Пользователь осуществляет регистрацию речевых данных с помощью мобильного приложения, выполняющего следующие этапы локальной обработки:

- предварительную фильтрацию и сглаживание;

- временную коррекцию (разделение длительных записей на сегменты до 30 с для удобства последующей обработки);
- преобразование стерео- в монозапись (включая понижение частоты дискретизации до 8000 Гц);
- преобразование в единый формат WAVE для хранения (включая изменение количества бит, используемых для передачи/обработки данных в единицу времени до 16 и использование импульсно-кодовой модуляции PCM без сжатия последовательно записанных значений амплитуды звука).



Рис. 1. Структура клиент-серверного решения системы обнаружения и классификации эмоциональных состояний человека по речевым сигналам

Подсистема обработки речевых сигналов разработана на языке программирования C++ и реализует математические модели, требующие необходимых вычислительных мощностей. Программное обеспечение функционирует на серверном устройстве системы, принимает исходный речевой сигнал и выполняет глобальную обработку, включающую в себя:

- подавление помех;
- сегментацию на информативные участки;

- определение информативных параметров;
- нормализацию значений информативных параметров.

Подсистема обработки речевых сигналов обеспечивает подготовку входных данных для подсистемы обнаружения и классификации критических эмоциональных состояний.

Подсистема обнаружения и классификации эмоциональных состояний человека разработана на языке программирования Python и реализует модели пятислойных нейронных сетей, обеспечивающие:

- гендерную классификацию;
- классификацию эмоциональных состояний;
- классификацию эмоций в рамках эмоциональных состояний.

Для эффективной программной реализации в Python использовались специализированные библиотеки SQLite3: NumPy, Tensorflow, модуль Keras библиотеки Tensorflow и модуль PyPlot библиотеки Matplotlib.

Подсистема управления взаимодействия процессами системы обнаружения и классификации эмоционального состояния человека разработана на языке программирования C++ и реализует:

- передачу речевого сигнала по сетевому протоколу передачи данных TCP/IP (Transmission Control Protocol/Internet Protocol), представленных в цифровом виде, посредством использования протокола шифрования данных HTTPS (HyperText Transfer Protocol Secure);

- передачу и отображение результат классификации;
- взаимодействие серверного и клиентского устройств.

Для аппаратной реализации системы обнаружения и классификации эмоциональных состояний человека по речевым сигналам в виде клиент-серверного решения выбраны и обоснованы следующие аппаратные средства:

- 1) материнская (системная) плата X99 D4X8 (производитель Atermiter) [10];
- 2) процессор XEON E5-2693v3 (производитель Intel) [11];
- 3) оперативная память DDR4 32GB (производитель Samsung) [12];
- 4) графический ускоритель GeForce RTX 3090 Ti (производитель NVIDIA) [13];
- 5) блок питания постоянного тока 2000W PSU 12V (производитель Taifast) [14];
- 6) система охлаждения;
- 7) система хранения данных уровня RAID 0.

Исследование системы

На рис. 2 представлена методика исследования системы обнаружения и классификации критических эмоциональных состояний человека по речевым сигналам в экстремальных условиях. Методика включает в себя два этапа: формирование выборки речевых сигналов и непосредственную классификацию эмоциональных состояний и эмоций человека.

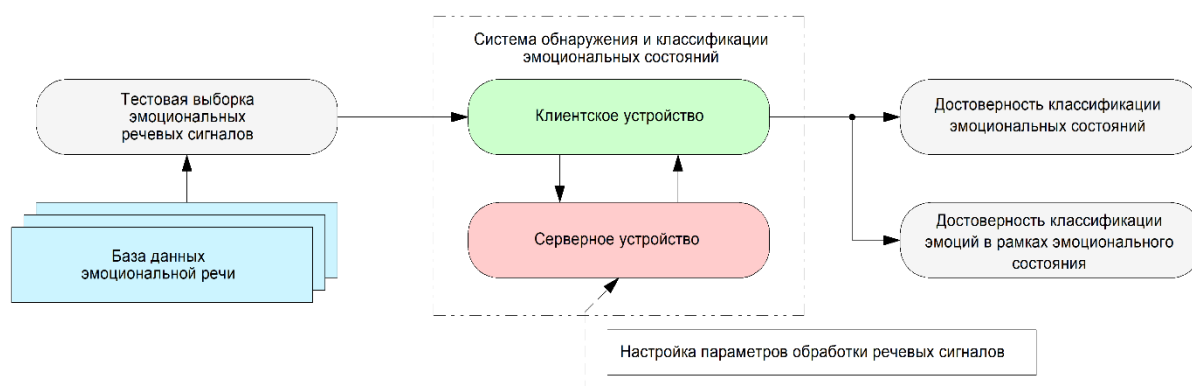


Рис. 2. Структура методики исследования системы обнаружения и классификации эмоциональных состояний человека по речевым сигналам

В рамках исследования предусмотрена возможность настройки параметров обработки речевых сигналов (использование этапов предварительной обработки и подавления помех). Суть

исследования заключается в оценке зависимости достоверности классификации разработанной системы от настройки параметров обработки.

Для оценки достоверности обнаружения и классификации критических отрицательных (страх, отвращение, гнев, печаль, удивление), нейтральных (нейтралитет) и положительных (удивление, радость) эмоциональных состояний человека использовались коэффициенты действительных и мнимых обнаружений, по которым определялись значения ошибок первого и второго рода.

На основе разработанной базы данных для исследования системы обнаружения и классификации эмоциональных состояний человека [15, 16] сформирована тестовая выборка эмоциональных речевых сигналов, включающая в себя 100 записей (50 записей – мужчин, 50 записей – женщин) словосочетаний и предложений из многосложных слов длительностью не более 10 с. Тестовая выборка включает в себя сигналы, зарегистрированные в условиях переживания эмоций человеком в реальных жизненных критических отрицательных, нейтральных и положительных эмоциональных состояниях.

В табл. 2–5 представлены усредненные результаты достоверности обнаружения и классификации эмоциональных состояний и эмоций для речевых сигналов тестовой выборки базы эмоциональной речи. В табл. 2 и 3 представлены результаты с использованием этапов предварительной обработки и подавления помех [17–20]. В табл. 4 и 5 представлены результаты без использования данных этапов.

Таблица 2

Достоверность классификации эмоциональных состояний
с предварительной обработкой и подавлением помех

Результат прогноза	Результат наблюдения		
	Отрицательное	Нейтральное	Положительное
Тестовая выборка базы эмоциональной речи			
Отрицательное	95,6	6,8	3,6
Нейтральное	7,5	94,7	5,8
Положительное	3,2	5,6	95,2

Таблица 3

Достоверность классификации эмоций с предварительной обработкой и подавлением помех

Результат прогноза	Результат наблюдения						
	Страх	Отвращение	Гнев	Печаль	Удивление	Нейтралитет	Радость
Тестовая выборка базы эмоциональной речи							
Страх	93,1	3,1	2,9	3,4	7,7	3,9	1,6
Отвращение	3,2	93,7	4,6	2,7	7,0	5,1	2,4
Гнев	2,7	4,3	95,6	2,4	4,8	1,8	8,2
Печаль	3,5	2,9	2,2	95,4	3,9	5,0	1,8
Удивление	8,1	6,8	4,4	3,8	94,7	6,6	3,7
Нейтралитет	4,1	5,4	1,9	4,5	6,5	95,2	2,4
Радость	1,7	2,1	6,7	1,5	4,6	2,3	95,5

Таблица 4

Достоверность классификации эмоциональных состояний
без предварительной обработки и подавлением помех

Результат прогноза	Результат наблюдения		
	Отрицательное	Нейтральное	Положительное
Тестовая выборка базы эмоциональной речи			
Отрицательное	92,7	7,1	3,8
Нейтральное	7,9	91,9	6,1
Положительное	3,4	5,9	92,3

Таблица 5

Достоверность классификации эмоций без предварительной обработки и подавлением помех

Результат прогноза	Результат наблюдения						
	Страх	Отвращение	Гнев	Печаль	Удивление	Нейтралитет	Радость
Тестовая выборка базы эмоциональной речи							
Страх	90,3	3,3	3,1	3,6	8,2	4,1	1,7
Отвращение	3,4	90,9	4,9	2,9	7,4	5,4	2,5
Гнев	2,9	4,6	92,7	2,5	5,1	1,9	8,7
Печаль	3,7	3,1	2,3	92,5	4,1	5,3	1,9
Удивление	8,6	7,2	4,7	4,0	91,9	7,0	3,9
Нейтралитет	4,3	5,7	2,0	4,8	6,9	92,3	2,5
Радость	1,8	2,2	7,1	1,6	4,9	2,4	92,6

В табл. 6 представлены значения достоверности обнаружения и классификации эмоциональных состояний человека по речевым сигналам разработанной системы в сравнении с аналогами.

Таблица 6

Достоверность обнаружения и классификации
эмоциональных состояний человека разработанной
системы в сравнении с аналогами

Наименование	Достоверность обнаружения и классификации
1. Система обнаружения и классификации эмоциональных состояний человека по речевым сигналам	94,7 – 95,6 % (для классификации эмоциональных состояний) 93,1 – 95,6 % (для классификации эмоций)
2. Система речевой аналитики «SVQA»	Не менее 90 %
3. Подсистема автоматической оценки качества «Smart Logger» многоканальной системы регистрации телефонных вызовов и речевых сообщений «НЕЗАБУДКА II»	Ошибка 1-го рода – не более 10 % Ошибка 2-го рода – не более 5 %
4. Профессиональная система диагностики голоса и речи «lingWAVES»	–
5. Технология анализа голоса «Sense»	Не менее 92 %
6. Технологии для анализа эмоций по голосу «Empath» / «Web Empath API» / «SDK Empath API» / «Beluga Box» / «Emo Value Generator»	–
7. Технология аналитики речи «SmartSpeech»	Не менее 90 %

Анализ результатов исследования

В соответствии с результатами в табл. 2–5 следует, что:

- достоверность классификации эмоциональных состояний и эмоций для речевых сигналов тестовой выборки базы эмоциональной речи составляет от 94,7 до 95,6 % и от 93,1 до 95,6 % соответственно;
- средние значения ошибок первого и второго родов классификации эмоциональных состояний и эмоций для речевых сигналов тестовой выборки базы эмоциональной речи с использованием этапов предварительной обработки и подавления помех в среднем на 3 и 6 % меньше соответственно, чем без использования;
- минимальное значение ошибки первого рода достигается при классификации противоположных отрицательных и положительных эмоциональных состояний 3,2 – 3,6 % и 3,3 – 3,7 % соответственно для речевых сигналов тестовой выборки базы эмоциональной речи;

– максимальное значение ошибки второго рода достигается при классификации эмоций страха и удивления 8,1 – 8,3 %, радости и гнева 8,2 – 8,4 % для речевых сигналов тестовой выборки базы эмоциональной речи;

– наилучшая достоверность классификации для речевых сигналов тестовой выборки базы эмоциональной речи достигается при классификации эмоций гнева, печали и радости 95,6, 95,4 и 95,5 % соответственно; наихудшая достоверность достигается при классификации страха – 93,1 %;

– достоверность обнаружения и классификации разработанной системы в среднем больше на 3,15 % для эмоциональных состояний и на 2,35 % для эмоций в сравнении с наилучшим аналогом – технологией «Sense».

Понижение средней достоверности классификации без использования этапов предварительной обработки и подавлением помех объясняется наличием физиологических шумов в естественной эмоциональной речи, возникающих вследствие нарушения работы органов речевого аппарата под воздействием эмоционального расстройства. Физиологические шумы вносят значительную погрешность, изменяя внутреннюю структуру полезного речевого сигнала.

Высокая достоверность классификации противоположных отрицательных и положительных эмоциональных состояний и эмоций гнева, печали и радости для речевых сигналов тестовой выборки базы эмоциональной речи объясняется высоким уровнем релевантности информативных параметров. Противоположная ситуация при классификации эмоций страха и удивления, радости и гнева – соответствующие информативные параметры имеют низкий уровень релевантности к данным эмоциям.

Заключение

В соответствии с анализом результатов исследования сделаны следующие краткие выводы:

1. Достоверность системы для классификации эмоциональных состояний составляет от 94,7 до 95,6 %, для классификации эмоций – от 93,1 до 95,6 %.

2. Время выполнения классификации составляет не более 30 с, при длительности анализируемой записи – не более 60 с.

3. Достоверность классификации обеспечивается высоким уровнем релевантности набора информативных параметров к эмоциональным состояниям и эмоциям.

4. Использование этапов предварительной обработки и подавления помех для фильтрации физиологических шумов в естественной эмоциональной речи обеспечивает повышение достоверности классификации критических эмоциональных состояний человека по речевым сигналам в экстремальных условиях.

В перспективе планируется провести дополнительные исследования быстродействия программно-аппаратной реализации системы обнаружения и классификации эмоциональных состояний человека по речевым сигналам.

Список литературы

1. Система речевой аналитики «SVQA». URL: <https://svqa.ru> (дата обращения: 20.02.2023).
2. Подсистема автоматической оценки качества «Smart Logger» многоканальной системы регистрации телефонных вызовов и речевых сообщений «НЕЗАБУДКА II». URL: <https://www.speechpro.ru> (дата обращения: 20.02.2023).
3. Профессиональная система диагностики голоса и речи «lingWAVES». URL: <https://www.wevosys.com> (дата обращения: 20.02.2023).
4. Технология анализа голоса «Sense». URL: <https://www.nemesysco.com> (дата обращения: 20.02.2023).
5. Технологии для анализа эмоций по голосу «Empath» / «Web Empath API» / «SDK Empath API» / «Beluga Vox» / «Emo Value Generator». URL: <https://www.webempath.com> (дата обращения: 20.02.2023).
6. Технология аналитики речи «SmartSpeech». URL: <https://sberdevices.ru/smartspeech> (дата обращения: 20.02.2023).

7. Alimuradov A. K., Tychkov A. Yu. Enhancing Estimation Accuracy of Psycho-Emotional State of Operators of Medical Exoskeletons as Intellectual Control Systems with High Degree of Responsibility // 2nd School on Dynamics of Complex Networks and Their Application in Intellectual Robotics (DCNAIR) (Saratov, Russia, 8–10 Oct., 2018). Saratov, 2018. P. 6–8.
8. Alimuradov A. K., Tychkov A. Yu. Enhancing estimation accuracy of psycho-emotional state of operators of medical exoskeletons as intellectual control systems with high degree of responsibility // Динамика сложных сетей и их применение в интеллектуальной робототехнике : сб. материалов II Междунар. школы-конференции молодых ученых (г. Саратов 8–10 октября 2018 г.). Саратов : Научная книга, 2018. С. 32–35.
9. Алимуратов А. К., Тычков А. Ю., Артемов И. И., Куликов М. Ю. Особенности построения систем голосового управления для медицинских экзоскелетов расширенного функционала // Качество. Инновации. Образование. 2016. № S2 (129). С. 66–72.
10. Материнская плата Atermiter X99 D4X8. URL: <https://intel.com> (дата обращения: 20.02.2023).
11. Процессор Intel XEON E5-2693v3. URL: <https://intel.com> (дата обращения: 20.02.2023).
12. Память оперативная Samsung DDR4 32GB. URL: <https://samsung.com> (дата обращения: 20.02.2023).
13. Графический ускоритель NVIDIA GeForce RTX 3090 Ti. URL: <https://nvidia.com> (дата обращения: 20.02.2023).
14. Блок питания постоянного тока Taifast Bitcorn ATX 2000W PSU 12V. URL: <https://taifast.com> (дата обращения: 20.02.2023).
15. Alimuradov A. K., Tychkov A. Yu., Gorbunov V. N. et al. Development of Natural Emotional Speech Database for Training Automatic Recognition Systems of Stressful Emotions in Human-Robot Interaction // 4th Scientific School on Dynamics of Complex Networks and their Application in Intellectual Robotics (DCNAIR) (Innopolis, Russia, 7–9 Sept., 2020). Innopolis, 2020. P. 11–16.
16. Свидетельство о государственной регистрации базы данных № 2020621102. Верифицированная база речевых сигналов, релевантно отражающих нарушения речевой функции человека, зарегистрированных в период стрессотравмирующих ситуаций (экстремальных условиях) / А. Ю. Тычков, А. К. Алимуратов, А. В. Агейкин, В. Н. Горбунов, Е. К. Рева, В. А. Межина // Программы для ЭВМ, базы данных, топологии интегральных микросхем ; заявл. 12.05.2020 ; опубл. 02.07.2020.
17. Алимуратов А. К., Тычков А. Ю., Чураков П. П., Дудников Д. С. Новый подход сегментации речевых сигналов на основе частотно-временного анализа // Измерение. Мониторинг. Управление. Контроль. 2022. № 4. С. 80–92.
18. Алимуратов А. К., Тычков А. Ю., Чураков П. П. [и др.]. Способ обработки речевых сигналов на основе метода декомпозиции на эмпирические моды // Измерение. Мониторинг. Управление. Контроль. 2022. № 2. С. 75–89.
19. Alimuradov A. K., Tychkov A. Yu., Churakov P. P. et al. Novel EMD-Based Technological Procedure for Speech Signal Processing // 24rd International Conference on Digital Signal Processing and its Applications (DSPA) (Moscow, Russia, 30 March – 1 April, 2022). Moscow, 2022. 6 p.
20. Alimuradov A. K., Tychkov A. Yu., Churakov P. P. et al. EMD-Based Method for Windowed Filtering of Speech Signals // 5th Scientific School Dynamics of Complex Networks and their Applications (DCNA) (Kaliningrad, Russia, 13–15 Sept., 2022). Kaliningrad, 2022. P. 6–9.

References

1. *Sistema rechevoy analitiki «SVQA» = Speech analytics system "SVQA"*. (In Russ.). Available at: <https://svqa.ru> (accessed 20.02.2023).
2. *Podsystema avtomaticheskoy otsenki kachestva «Smart Logger» mnogokanal'noy sistemy registratsii telefonnykh vyzovov i rechevykh soobshcheniy «NEZABUDKA II» = Subsystem of automatic quality assessment "Smart Logger" of the multichannel system of registration of telephone calls and voice messages "FORGET-ME-NOT II"*. (In Russ.). Available at: <https://www.speechpro.ru> (accessed 20.02.2023).
3. *Professional'naya sistema diagnostiki golosa i rechi «lingWAVES» = Professional voice and speech diagnostics system "lingWAVES"*. (In Russ.). Available at: <https://www.wevosys.com> (accessed 20.02.2023).
4. *Tekhnologiya analiza golosa «Sense» = Voice analysis technology "Sense"*. (In Russ.). Available at: <https://www.nemesysco.com> (accessed 20.02.2023).
5. *Tekhnologii dlya analiza emotsiy po golosu «Empath» = Technologies for analyzing emotions by voice "Empath" / «Web Empath API» / «SDK Empath API» / «Beluga Box» / «Emo Value Generator»*. Available at: <https://www.webempath.com> (accessed 20.02.2023).
6. *Tekhnologiya analitiki rechi «SmartSpeech» = SmartSpeech speech analytics technology*. (In Russ.). Available at: <https://sberdevices.ru/smartspeech> (accessed 20.02.2023).

7. Alimuradov A.K., Tychkov A.Yu. Enhancing Estimation Accuracy of Psycho-Emotional State of Operators of Medical Exoskeletons as Intellectual Control Systems with High Degree of Responsibility. *2nd School on Dynamics of Complex Networks and Their Application in Intellectual Robotics (DCNAIR) (Saratov, Russia, 8–10 Oct., 2018)*. Saratov, 2018:6–8.
8. Alimuradov A.K., Tychkov A.Yu. Enhancing estimation accuracy of psycho-emotional state of operators of medical exoskeletons as intellectual control systems with high degree of responsibility. *Dinamika slozhnykh setey i ikh primeneniye v intellektual'noy robototekhnike: sb. materialov II Mezhdunar. shkoly-konferentsii molodykh uchenykh (g. Saratov 8–10 oktyabrya 2018 g.) = Dynamics of complex networks and their application in intelligent robotics : collection of materials II International. schools-conferences of young scientists (Saratov, October 8-10, 2018)*. Saratov: Nauchnaya kniga, 2018: 32–35.
9. Alimuradov A.K., Tychkov A.Yu., Artemov I.I., Kulikov M.Yu. Features of the construction of voice control systems for medical exoskeletons of extended functionality. *Kachestvo. Innovatsii. Obrazovanie = Quality. Innovation. Education*. 2016;(S2):66–72. (In Russ.)
10. *Materinskaya plata Atermiter X99 D4X8 = Atermiter X99 D4X8 motherboard*. (In Russ.). Available at: <https://intel.com> (accessed 20.02.2023).
11. *Protsessor Intel XEON E5-2693v3 = Intel XEON E5-2693v3 processor*. (In Russ.). Available at: <https://intel.com> (accessed 20.02.2023).
12. *Pamyat' operativnaya Samsung DDR4 32GB = Samsung DDR4 32GB RAM*. (In Russ.). Available at: <https://samsung.com> (accessed 20.02.2023).
13. *Graficheskiy uskoritel' NVIDIA GeForce RTX 3090 Ti = NVIDIA GeForce RTX 3090 Ti Graphics Accelerator*. (In Russ.). Available at: <https://nvidia.com> (accessed 20.02.2023).
14. *Blok pitaniya postoyannogo toka Taifast Bitcorn ATX 2000W PSU 12V = Taifast Bitcorn ATX 2000W PSU 12V DC power supply*. Available at: <https://taifast.com> (accessed 20.02.2023).
15. Alimuradov A.K., Tychkov A.Yu., Gorbunov V.N. et al. Development of Natural Emotional Speech Database for Training Automatic Recognition Systems of Stressful Emotions in Human-Robot Interaction. *4th Scientific School on Dynamics of Complex Networks and their Application in Intellectual Robotics (DCNAIR) (Innopolis, Russia, 7–9 Sept., 2020)*. Innopolis, 2020:11–16.
16. Certificate of state registration of the database No. 2020621102. The verified database of speech signals, relevant to reflect violations of human speech function, registered during stress-traumatic situations (extreme conditions) / A.Y. Tychkov, A.K. Alimudov, A.V. Ageikin, V.N. Gorbunov, E.K. Reva, V.A. Mezhdina // Computer programs, databases, topologies integrated circuits ; application. 12.05.2020 ; publ. 02.07.2020. (In Russ.)
17. Alimuradov A.K., Tychkov A.Yu., Churakov P.P., Dudnikov D.S. A new approach to segmentation of speech signals based on time-frequency analysis. *Izmerenie. Monitoring. Upravlenie. Kontrol' = Measurement. Monitoring. Management. Control*. 2022;(4):80–92. (In Russ.)
18. Alimuradov A.K., Tychkov A.Yu., Churakov P.P. et al. Method of processing speech signals based on the method of decomposition into empirical modes. *Izmerenie. Monitoring. Upravlenie. Kontrol' = Measurement. Monitoring. Management. Control*. 2022;(2):75–89. (In Russ.)
19. Alimuradov A.K., Tychkov A.Yu., Churakov P.P. et al. Novel EMD-Based Technological Procedure for Speech Signal Processing. *24rd International Conference on Digital Signal Processing and its Applications (DSPA) (Moscow, Russia, 30 March – 1 April, 2022)*. Moscow, 2022:6.
20. Alimuradov A.K., Tychkov A.Yu., Churakov P.P. et al. EMD-Based Method for Windowed Filtering of Speech Signals. *5th Scientific School Dynamics of Complex Networks and their Applications (DCNA) (Kaliningrad, Russia, 13–15 Sept., 2022)*. Kaliningrad, 2022:6–9.

Информация об авторах / Information about the authors

Алан Казанферович Алимуратов

кандидат технических наук,
директор студенческого научно-
производственного бизнес-инкубатора,
Пензенский государственный университет
(Россия, г. Пенза, ул. Красная, 40)
E-mail: alansapfir@yandex.ru

Alan K. Alimuradov

Candidate of technical sciences,
director of student research and production
business incubator,
Penza State University
(40 Krasnaya street, Penza, Russia)

Автор заявляет об отсутствии конфликта интересов /
The author declares no conflicts of interests.

Поступила в редакцию/Received 27.02.2023

Поступила после рецензирования/Revised 24.03.2023

Принята к публикации/Accepted 24.04.2023