



ских расстройств / А.К. Алимуратов, А.Ю. Тычков, П.П. Чураков, А.В. Агейкин // Известия высших учебных заведений. Поволжский регион. Технические науки. - 2017. - № 3 (47). - С. 4 - 25.

8. Schuller B.W., Batliner A.M. Computational Paralinguistics: Emotion, Affect and Personality in Speech and Language Processing. New York: Wiley, 2013, 344 p.

А.К. Алимуратов, А.Ю. Тычков, П.П. Чураков, А.В. Агейкин

ОЦЕНКА ПСИХОЭМОЦИОНАЛЬНОГО СОСТОЯНИЯ НА ОСНОВЕ АНАЛИЗА РЕЧЕВЫХ СИГНАЛОВ: СОВРЕМЕННОЕ СОСТОЯНИЕ, ПРОБЛЕМЫ И ПЕРСПЕКТИВЫ

(Пензенский государственный университет)

Введение

Оценка психоэмоционального состояния населения на сегодняшний день является социально-значимой проблемой для каждого государства, поскольку напрямую связана с формированием здорового образа жизни населения и профилактикой психических заболеваний. По данным ВОЗ современные социально-значимые заболевания, являющиеся главной причиной временной нетрудоспособности, инвалидности и смертности населения, напрямую связаны с психоэмоциональным состоянием населения.

В статье кратко представлен обзор современного состояния области оценки психоэмоционального состояния человека на основе речевых сигналов; выявлены проблемы и определены перспективы развития данной области. Работа выполнена при финансовой поддержке Совета по грантам Президента РФ в рамках проекта № СП-246.2018.5 и является продолжением ранее опубликованных трудов авторов [1, 2].

Обзорный анализ [3 - 5] выявил, что оценка психоэмоционального состояния человека на основе анализа речи сводится к выполнению трех ключевых задач: выбор/создание базы исследуемых речевых сигналов, выделение информативных параметров и построение алгоритма классификации психоэмоциональных состояний.

База данных речевых сигналов

Проведен обзор 64 баз данных эмоционально окрашенных речевых сигналов [5]. Базы были систематизированы по следующим параметрам: язык; количество и профессия испытуемых; физиологические особенности; виды психоэмоциональных состояний (радость, нейтралитет, скука, печаль, раздражение, страх, отвращение и гнев); условия их воспроизведения (вынужденные, смоделированные и естественные). Вынужденные и смоделированные эмоции, как правило, воспроизводятся профессиональными актерами. Речь, произносимая профессионалами, является самой эмоциональной окрашенной, так как эмоции актеров имеют большой диапазон амплитуды и продолжительности. Но основ-



ным их недостатком является тот факт, что они не похожи на естественно выраженные психоэмоциональные состояния. Под естественно выраженными психоэмоциональными состояниями, подразумеваются эмоции, воспроизводимые в результате реакции на естественные ситуации. На самом деле, в реальной жизни, диалог между людьми может содержать массу эмоций и соотносить одно психоэмоциональное состояние с одной ситуацией затруднительно.

Информативные параметры

Речь представляет собой нестационарный акустический сигнал сложной формы, амплитудные и частотные характеристики которого быстро изменяются во времени. Речь – один из самых сложных приобретаемых навыков органов речевого аппарата и очень чувствительна к нарушениям работы скоплений нервных окончаний, участвующих в формировании речевых сигналов. Вид и степень выраженности психоэмоционального состояния кодируются в определенные акустические и лингвистические параметры речи [3, 4].

Акустические параметры. Данные параметры в большей степени характеризуют не сами психоэмоциональные состояния, а их поведение в течение времени. Идея использования акустических параметров для оценки психоэмоциональных состояний основывается на том факте, что люди используют вольные или невольные акустические вариации, чтобы подчеркнуть значимость отдельных элементов речи [3]. К акустическим параметрам относятся: интонация и интенсивность речи, коэффициенты предсказания и кепстральные коэффициенты, спектр и форманты, гармоничность и искажение речи, и др.

Лингвистические параметры. Слова или грамматические конструкции, используемые в речи, играют немаловажную роль в отражении психоэмоционального состояния. Выделение лексем может быть достигнуто путем преобразования текста в классы слов. Первый класс состоит из лексем, называется морфологическим и он группирует слова по их основам. Второй класс состоит из частей речи, таких как существительные, глаголы, прилагательные. Наконец, семантические единицы, представленные лексемами, могут быть сгруппированы в смысловые понятия высшего порядка, например, положительные или отрицательные выражения. [4, 5].

Функционалы. Данные вид информативных признаков представляет ряд математических операторов, применяемых к акустическим параметрам для получения векторов признаков одинаковой размерности. Функционалы обеспечивают своего рода нормализацию по времени, например можно получить один вектор параметров для одного слова с постоянным числом элементов [4, 5]. К функционалам относятся: среднее и стандартное отклонения, асимметрия, значения экстремумов, диапазон значений, моменты высшего порядка, пики, сегменты и др.

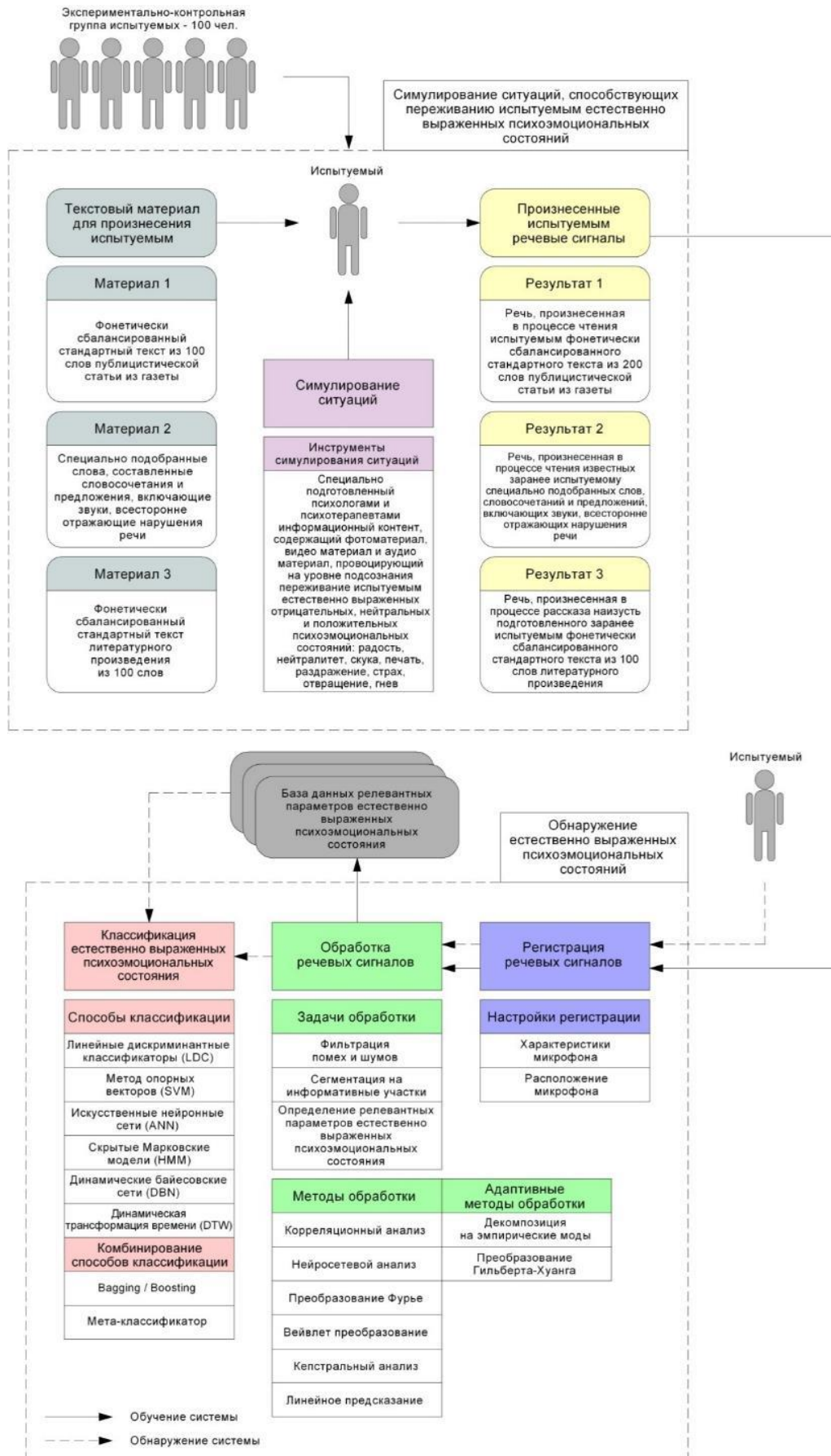


Рисунок 1 – Структура перспективных исследований в области оценки психоэмоционального состояния



Алгоритмы классификации психоэмоциональных состояний

Линейные дискриминантные классификаторы (*Linear Discriminant Classifiers, LDCs*) и классификаторы k -ближайших соседей (*k-Nearest Neighbour, kNN*) пользуются широкой популярностью, так как эффективны для смоделированных и выраженных психоэмоциональных состояний, но малоприспособлены для большого набора параметров.

Также хорошо известен метод опорных векторов (*Support Vector Machines, SVM*), который является продолжением метода *LDCs* и демонстрирует хорошие результаты для большого количества информативных параметров. Наиболее часто используемыми нелинейными дискриминантными классификаторами, являются искусственные нейронные сети (*Artificial Neural Networks, ANNs*). Для метода *ANNs* необходимо большее количество данных, и поэтому они редко используются для смоделированных психоэмоциональных состояний, и еще меньше для естественных эмоций. Динамические классификаторы, такие как динамические байесовские сети (*Dynamic Bayesian Networks*), скрытые марковские модели (*Hidden Markov Models*) или простая динамическая трансформация времени (*Dynamic Time Warp*) обеспечивают меньшую точность классификации, чем статические. В работе [5] отмечается, что сочетание статической и динамической классификации может повысить точность обнаружения эмоций. Популярными подходами, объединяющими классификаторы, являются *Bagging* и *Boosting*.

Проблемы и перспективы развития

В соответствии с проведенным обзором, отметим ряд нерешенных проблем в данной области:

- создание базы данных речевых сигналов с естественными эмоциями, зарегистрированными в реальных условиях;
- создание модели информативных параметров в условиях проявления разных видов психоэмоциональных состояний - не установлена четкая взаимосвязь вида эмоционального состояния и объективного набора релевантных параметров речевых сигналов;
- оптимизация задач регистрации естественной эмоциональной речи, определения релевантного набора параметров и классификации эмоций в реальном времени для системы оценки психоэмоционального состояния.

На рисунке 1 структурно представлены перспективы развития данной области.

Литература

1. Алимуратов А.К. Способ определения кепстральных маркеров речевых сигналов при психогенных расстройствах / А.К. Алимуратов, А.Ю. Тычков, А.П. Зарецкий, А.П. Кулешов // Труды МФТИ. - 2017. - Т. 9, № 4. - С. 201 - 214.
2. Тычков А.Ю. Современные представления о специфике и проблемах диагностики пограничных психических расстройств / А.Ю. Тычков, А.В. Агейкин, А.К. Алимуратов, В.Б. Калистратов, С.Ю. Митрошина // Психическое здоровье. - 2017. - Т. 15. - № 5. - С. 69 - 75.



3. Schuller B.W., Batliner A.M. Computational Paralinguistics: Emotion, Affect and Personality in Speech and Language Processing. New York: Wiley, 2013, 344 p.

4. Schuller B. Recognising realistic emotions and affect in speech: State of the art and lessons learnt from the first challenge / B. Schuller, A. Batliner, S. Steidl, D. Seppi // Speech Communication. - 2011. - V. 53 (9-10). - P. 1062 - 1087.

5. Ververidis D. Emotional speech recognition: Resources, features, and methods / D. Ververidis, C. Kotropoulos // Speech Communication. - 2006. - V. 48 (9). - P. 1162 - 1181.

Б.Г. Алмухамедова¹, Д.А. Алиева², Б.Д. Юсупов², О.И Рузибоев³

НЕОБХОДИМОСТЬ СОЗДАНИЯ ИНФОРМАЦИОННОЙ СИСТЕМЫ ДЛЯ РАННЕГО ВЫЯВЛЕНИЯ РАКА ШЕЙКИ МАТКИ В УЗБЕКИСТАНЕ

¹Ташкентский городской филиал РСНПМЦОиР МЗ РУз,

²РСНПМЦОиР МЗ РУз,

³Ташкентский университет информационных технологий имени ал-Хоразми)

This article discusses the possibility of creating a computer program for the formation of a database of risk groups among women under examination for the early detection of cervical cancer and previous pathology. The established program will contribute to the formation and storage of the database, as well as the ability to monitor and timely control.

Злокачественные новообразования (ЗН) у женщин, такие как рак молочной железы (РМЖ), рак шейки матки (РШМ), рак яичников (РЯ), являются причиной преждевременной смерти сотен тысяч женщин. В структуре злокачественных новообразований (ЗН) женщин Узбекистана РШМ занимает второе место после рака молочной железы (РМЖ), при этом на его долю приходится 12,4%.

Стандартизированный показатель заболеваемости РШМ по республике составляет 5,5 на 100 тыс. населения, «грубый» – 4,7, [1] что выдвигает в первую очередь одну из глобальных проблем – решение вопроса раннего выявления фоновых и предраковых заболеваний, а также ранних стадий заболевания. При этом оказание своевременной помощи и проведение адекватного специфического лечения позволит снизить смертность и улучшить показатели выживаемости больных, тем самым обеспечивая высокий жизненный потенциал и качество жизни.

Существенное инвестирование в разработку программ, направленных на профилактику и лечение предраковых и раковых заболеваний у женщин, таких как РМЖ и РШМ значительно улучшили ситуацию и соответственно позволили снизить смертность от этих онкологических заболеваний в странах с высоким уровнем доходов[4]. Разработка и проведение широкомасштабных программ является затратным и требует больших финансовых вложений. В результате