

ОТЗЫВ

на автореферат диссертации СТЕФАНИДИ Антона Федоровича на тему: «Исследование мультимодальных алгоритмов биометрической идентификации на основе методов цифровой обработки речевых сигналов и изображений», представленную на соискание ученой степени кандидата технических наук по специальности 2.2.13 – Радиотехника, в том числе системы и устройства телевидения (Федеральное государственное бюджетное образовательное учреждение высшего образования «Владимирский государственный университет имени Александра Григорьевича и Николая Григорьевича Столетовых», 600000, г. Владимир, ул. Горького, д.87 ВлГУ)

В настоящее время ни у кого не вызывает сомнения тот факт, что современное общество поступательно переходит в информационную стадию развития, что, в свою очередь, предполагает цифровую трансформацию за счёт совершенствования информационных и телекоммуникационных технологий. Под возрастающим влиянием этих процессов происходит коренная перестройка всего спектра производственных и социальных отношений, формируются новые модели общественного развития. В ходе внедрения современных информационных и телекоммуникационных технологий важная роль отводится надёжности функционирования различных систем идентификации (в т.ч. биометрической) пользователей при их взаимодействии с разнообразными информационными системами и устройствами. Таким образом, проводимое в рамках диссертации исследование вопросов совершенствования систем биометрической идентификации на основе методов цифровой обработки речевых сигналов и изображений представляется актуальным.

Одним из перспективных направлений развития систем биометрической идентификации личности является совершенствование алгоритмов распознавания речи и образов на основе многофакторного статистического анализа. При этом очевидно, что исследование вопросов построения надёжных алгоритмов работы, обеспечивающих функционирование систем идентификации и аутентификации личности на основе биометрических параметров человека, таких как индивидуальные особенности речи и черты лица, имеет высокое практическое значение.

Представленный к рассмотрению автореферат диссертации соискателя Стефаниди А.Ф. позволяет сделать вывод, что заявленная соискателем тема

исследования соответствует формуле специальности 2.2.13 – Радиотехника, в том числе системы и устройства телевидения в части пунктов 4-6 паспорта специальности.

Объектом исследования автора являются алгоритмы биометрической идентификации, применяемые в системах прикладного телевидения и радиотехнических системах обработки и анализа цифровых сигналов.

Предметом исследования определены структуры нейросетевых алгоритмов биометрической идентификации на основе анализа речевых сигналов и цифровых изображений лиц с целью повышения достоверности распознавания личностей.

Целью диссертационного исследования заявлено повышение точности систем идентификации личности путём разработки нейросетевых алгоритмов анализа речевых сигналов и изображений лиц.

Для достижения поставленной цели автором сформулированы следующие задачи исследования:

разработка комбинированного детектора голосовой активности;

разработка нейросетевых алгоритмов идентификации личности на основе анализа речевых сигналов и изображений лиц;

усовершенствование работы алгоритмов идентификации личности в условиях действия шумов и помех в речевых сигналах и наличия медицинской маски на изображениях лиц;

разработка мультимодальных алгоритмов идентификации личности на основе комбинированного анализа речевых сигналов и изображений лиц.

В качестве методов диссертационного исследования, помимо общих системных, автор использует цифровую обработку сигналов и изображений, спектральный анализ, распознавание образов в нейронных сетях. Для машинной реализации алгоритмов обучения нейронной сети и распознавания образов обученной сетью применялись программные средства языка программирования Python, а также библиотеки языка Python для обучения нейронных сетей TensorFlow и Keras.

К научным достижениям диссертационных исследований автора могут быть отнесены результаты разработок:

архитектуры комбинированного детектора голосовой активности для выделения речевых фрагментов аудиозаписей на основе алгоритма решающих деревьев;

робастного алгоритма цифровой обработки аудиосигналов голосовой биометрии на основе предложенной автором структуры нейросети,

обеспечивающий низкую деградацию качества в условиях действия шумов и помех;

робастного алгоритма обработки лицевой биометрии на основе реализации сверточной схемы нейронных сетей, обеспечивающей низкую деградацию качества в условиях наличия медицинской маски;

мультимодальные алгоритмы идентификации личности для комплексной обработки голосовой и лицевой биометрий для распознавания образа и окончательного принятия решения по идентификации личности.

Теоретическая значимость диссертационного исследования заключается в построении с применением нейросетей методики комплексного решения двух разновидностей прикладных задач распознавания образов (голосовой и лицевой биометрий), что в дальнейшем может служить шагом к совершенствованию методологического подхода систем искусственного интеллекта.

Практическая значимость результатов, полученных в диссертации, заключается в том, что применительно к специфике решаемых автором прикладных задач идентификации и аутентификации личности проведены оценки емкостных и временных сложностей существующих унимодальных алгоритмов распознавания образов, а также достигаемых точностей распознавания личностей предложенными алгоритмами.

Теоретические положения диссертационной работы могут быть использованы как в учебном процессе при подготовке специалистов соответствующих направлений, так и при обосновании требований к разнообразным системам распознавания образов при их проектировании, развертывании и эксплуатации. Ценность исследования состоит также в том, что предлагаемый подход, реализованный в разработанных комплексных алгоритмах взаимодействия нейросетей идентификации личностей может быть сравнительно легко распространен на другие задачи семейства распознавания образов.

Достоверность предложенных автором мультимодальных алгоритмов подтверждена практической проверкой составленными на языке Python программами, реализующими эти алгоритмы. Для проверки автором были подготовлены большой объем данных для тестирования алгоритмов на ЭВМ и сравнения результатов распознавания образов с результатами обработки тех же данных унимодальными алгоритмами. Обоснованность результатов выполненного научного исследования обеспечивается целостным, комплексным подходом к разработке методики его проведения,

адекватностью выбранных методов цифровой обработки сигналов, подтверждением результатов численными экспериментами, научной апробацией основных идей и результатов на многочисленных научных конференциях, включая международные.

Автореферат диссертации написан достаточно грамотным в техническом отношении языком и вполне отражает суть диссертационной работы. Общая характеристика исследования, основное содержание работы, теоретические и практические выводы в автореферате диссертации изложены в логической взаимосвязи друг с другом и сбалансированы по своему содержанию. Имеют место некоторое огрехи стилистического и пунктуационного характера.

Содержание автореферата и приведенных публикаций дает вполне ясное понимание о проделанной работе, отражает основные идеи, достижения и выводы диссертации.

Приведенные диссертантом выводы и рекомендации соответствуют цели и задачам исследования, являются убедительными и достоверными, результаты подтверждены вычислительными экспериментами.

В качестве замечаний и рекомендаций к автореферату можно указать следующее:

1. На наш взгляд следовало бы более детально структурировать задачу идентификации личностей, отдельно рассматривая две её части: задачу таксономии (классификации), т.е. выявления в ходе обучения количества и характерных значений параметров, присущих классам схожести личностей друг с другом (классов эквивалентности элементов) и задачу распознавания тестируемой личности, т.е. установления принадлежности образа тому или иному классу эквивалентности (таксону). Это позволило бы более корректно устанавливать границы применимости алгоритмов. Так, если в результате обучения каждый таксон содержит ровно один элемент различимый образ из исходной выборки, то алгоритм полностью пригоден для решения задач СОРМ, и в этом случае большие выбросы отсеивать нельзя. В противном случае задача опознавания конкретной личности может быть решена лишь с точностью до утверждения «с наибольшей вероятностью личность похожа на такого-то» (или, более формально, «с наибольшей вероятностью распознаваемая личность относится к такому-то классу»);

2. Применение автором предварительной обработки выборки входных голосовых данных и отбрасывания их из генеральной выборки звуковых данных с одной стороны не лишена практического смысла, но с другой

стороны препятствует применению такого подхода для распознавания характерных длительностей пауз в речи индивида;

3. С формальной точки зрения эффект мультимодальности обработки мог бы быть достигнут объединением исходных параметров аудиозаписей и изображений индивидов в единое исходное пространство рабочих параметров. Автор же предпочёл объединить процессы обучения и непосредственного распознавания личностей на уровне результатов отдельной обработки изображений и аудиозаписей, что на наш взгляд приводит к игнорированию корреляции между спецификой речи и изображениями личностей. На наш взгляд автору не следовало бы бездоказательно отбрасывать эффект такой взаимной корреляции, тем более, что одним из аспектов новизны исследований является комплексность использования данных для распознавания личностей;

4. Прием по растягиванию изображения, предложенный автором, в случае наличия медицинской маски на лице человека, изложен не вполне определенно. Так искусственное растяжение верхней части лица человека с использованием одного и того же исходного изображения, видимо, должно быть полностью компенсировано за счёт последующего цифрового совмещения фрагмента с полным изображением (при этом при безошибочной цифровой обработке разрешение изображения фактически не изменится), а если же верхняя часть лица снимается отдельно с большим вертикальным разрешением, не совсем понятно, по какой причине автором не рассматривается возможность получения снимка с большим разрешением и для нижней части лица с выполнением последующего цифрового совмещения нижней и верхней части лица по вертикали;

5. На наш взгляд, автору, наряду со сравнением мультимодальных алгоритмов с унимодальными аналогами по временным сложностям и количеству значимых параметров результатов таксономии, следовало бы акцентировать внимание на анализе и причинах различий в перечнях значимых параметров после обучения для сравниваемых алгоритмов.

Однако, указанные замечания и рекомендации касаются в основном методологии построения систем искусственного интеллекта, которая непосредственно не входит в сферу диссертационных исследований автора. Теория распознавания образов использована автором исключительно для демонстрации работоспособности предложенных им решений задач диссертационного исследования. Указанное выше не снижает общей высокой оценки представленного автореферата диссертационной работы.

Вывод: содержание автореферата свидетельствует, что диссертация Стефаниди А.Ф. является самостоятельно выполненной, законченной научно-квалификационной работой, имеющей теоретическое и практическое значение в области радиотехники и отвечает требованиям «Положения о порядке присуждения ученых степеней», предъявляемым к кандидатским диссертациям, а её автор заслуживает присуждения ученой степени кандидата технических наук по специальности 2.2.13. – Радиотехника, в том числе системы и устройства телевидения.

Директор Центра исследования интернета вещей ФГБУ НИИР,
доктор военных наук доцент

П.И. Антонович

«9» августа 2022 г.

Подпись доктора военных наук доцента Антоновича Павла Игоревича, директора Центра исследования интернета вещей ФГБУ «Ордена Трудового Красного Знамени Российский научно-исследовательский институт радио имени М.И. Кривошеева» подтверждаю

Ученый секретарь ФГБУ НИИР
кандидат технических наук, СНС

М.М. Ступницкий

«9» августа 2022 г.

