

ОТЗЫВ

официального оппонента на диссертацию

Рахманенко Ивана Андреевича

**на тему «Алгоритмы и программные средства верификации
диктора по произвольной фразе»**

**по специальности 05.13.17 – «Теоретические основы информатики»
на соискание ученой степени кандидата технических наук**

Актуальность избранной темы

В последнее время были достигнуты значительные успехи в верификации личности по голосу и другим модальностям. Однако исчерпаны далеко не все резервы по повышению надежности биометрической верификации личности. Уникальность голосовой биометрии состоит в том, что это единственная биометрическая модальность, которая позволяет идентифицировать человека по телефону. Это важно, например, при удаленном доступе к различным услугам, при криминалистической идентификации, где единственным доказательством является запись телефонного разговора подозреваемого. Голосовая верификация не требует применения специализированного дорогостоящего оборудования. Все, что необходимо – обычный микрофон. При этом по уровню надежности голосовая биометрия не уступает, а по некоторым характеристикам превосходит характеристики других систем биометрической идентификации.

Оппонируемая диссертационная работа Рахманенко И.А. направлена на исследование речевых признаков, используемых для верификации диктора, с целью повышения точности методов верификации диктора по произвольной фразе. Одним из способов решения этой задачи является применение робастных признаков, применение признаков различной модальности. Таким образом, можно заключить, что выбранная тема и направленность исследования научной работы являются актуальными.

Степень обоснованности научных положений, выводов и рекомендаций, сформулированных в диссертации

Обоснованность результатов диссертационного исследования подтверждена широким использованием научных работ отечественных и зарубежных авторов в области обработки речи, идентификации и верификации диктора. Автором изучены и критически проанализированы достижения и положения других авторов по вопросам извлечения и применения речевых признаков для верификации диктора по произвольной фразе. Список содержит 139 литературных источников. Достаточный уровень объективности и адекватности выводов 1) базируется на теоретических и экспериментальных исследованиях, выполненных с использованием современных методов и средств; 2) подтверждается использованием в работе последних публикаций в периодических изданиях по исследуемой тематике.

Достоверность и новизна, полученных результатов

Достоверность полученных результатов обеспечивается строгостью используемого математического аппарата, адекватностью и достоверностью алгоритмов, установленной путем сравнения реакции алгоритмов и известных аналогов на тестовые данные, большим количеством экспериментальных данных, подтверждающих теоретические результаты, а также внедрением разработанного комплекса алгоритмов в практику.

Новизна проведенных исследований заключается в разработке оригинального алгоритма верификации диктора, отличающегося от существующих применением речевых признаков, полученных с помощью жадного алгоритма отбора признаков; разработке алгоритма генерации признаков, основанный на применении сверточной глубокой сети доверия; разработке нового гибридного алгоритма верификации диктора по произвольной фразе на основе ансамбля классификаторов.

Теоретическая и практическая значимость полученных автором результатов

Теоретическая ценность работы заключается в развитии алгоритмов и методов извлечения признаков из данных, создании новых алгоритмов в области обработки речи. Алгоритм генерации признаков, основанный на применении сверточной глубокой сети доверия, может использоваться не только для выделения признаков из речевых данных, но и для выделения признаков из изображений. Результаты работы могут использоваться в таких задачах как идентификация пола диктора и распознавание речи.

Практическая значимость представляется разработанным программным средством, которое может быть применено для верификации диктора по произвольной фразе. Полученное автором программное средство используются в АО «ОЭЗ ТВТ «Томск».

Созданные алгоритмы и программное средство использованы в рамках мероприятия 1.3 ФЦП «Исследования и разработки по приоритетным направлениям развития научно-технологического комплекса России на 2014—2020 годы» (соглашение о предоставлении субсидии № 14.577.21.0172 от 27 октября 2015 г.; уникальный идентификатор RFMEFI57715X0172), в рамках выполнения базовой части государственного задания Минобрнауки России, проект 8.9628.2017/8.9 на базе лаборатории медико-биологических исследований (ЛМБИ) ТУСУР.

Оценка содержания диссертации, её завершенность

Основные результаты диссертации изложены в трех главах, в которых достигается цель по разработке и исследованию алгоритмов и речевых признаков верификации диктора по произвольной фразе, повышающих точность верификации диктора по сравнению с известными подходами, методами и алгоритмами.

Во введении обоснована актуальность темы диссертации, сформулированы цели и задачи работы, указаны научная новизна и практическая значимость полученных результатов, представлена структура диссертации.

В главе 1 автором была произведена постановка задачи верификации диктора по произвольной фразе, проведен подробный анализ состояния проблемы исследования. Были рассмотрены основные речевые признаки, используемые в системах верификации диктора, основные современные методы верификации диктора по произвольной фразе, а именно: метод верификации, основанный на применении гауссовых смесей, метод верификации, основанный на факторном анализе, методы верификации с применением глубоких нейронных сетей.

В главе 2 автором, основываясь на проведенном литературном обзоре, для повышения точности методов верификации диктора по произвольной фразе были предложены алгоритмы верификации с применением различных речевых признаков. В первой части главы подробно рассмотрен алгоритм верификации на базе гауссовых смесей и универсальной фоновой модели, принятый в качестве базового. Произведено экспериментальное исследование и оценена точность системы верификации диктора с применением различных наборов признаков.

Во второй части главы было произведено исследование речевых признаков с применением генетического алгоритма и жадного алгоритма добавления-удаления. Данные алгоритмы были использованы для определения набора признаков, дающего наибольшую точность системы верификации диктора по произвольной фразе. При проведении эксперимента на рассмотренном в работе речевом корпусе, наилучшие результаты по точности показал алгоритм жадного добавления-удаления. В набор признаков, полученный данным алгоритмом, входят такие признаки как мел-кепстральные коэффициенты, дельта-коэффициенты, коэффициент линейного предсказания, линейная спектральная пара и вероятность вокализации. Положительным моментом полученного алгоритма можно назвать относительно небольшое количество используемых признаков при существенном увеличении точности системы. Также был проведен эксперимент с применением речевого корпуса CHAINS, на котором система верификации показала точность 100 %, что можно также оценить положительно.

В третьей части главы представлен алгоритм генерации признаков, основанный на применении сверточной глубокой сети доверия. Показано экспериментальное исследование с применением признаков, извлеченных с помощью сверточной глубокой сети доверия, в котором в качестве классификатора была использована гауссова смесь. Можно отметить то, что точность верификации диктора с применением предложенных признаков оказалась ниже, чем с применением общеизвестных признаков.

В четвертой части главы представлен гибридный алгоритм верификации диктора по произвольной фразе на основе ансамбля классификаторов, в котором были использованы результаты, полученные во второй и третьей частях главы. Для решения задачи верификации диктора по произвольной фразе автор предлагает новый алгоритм, использующий как гауссовы смеси, так и такие классификаторы как машина опорных векторов, классификатор, основанный на линейном дискриминантном анализе, классификатор AdaBoost M1. Результаты

эксперимента показали, что применение высокоуровневых признаков, извлекаемых с помощью третьего слоя сверточной глубокой сети доверия, вносит определенный вклад в повышение точности алгоритма верификации диктора.

В главе 3 рассмотрено разработанное программное средство для верификации диктора по произвольной фразе, применение предложенного программного средства для решения задачи контроля доступа к рабочим компьютерам сотрудников. С помощью метода анализа иерархий было произведено сравнение представленного программного средства с аналогами, которое показало преимущество разработки по точности верификации диктора.

В приложениях приведены акты о внедрении результатов диссертационной работы: в АО «Особая экономическая зона технико-внедренческого типа «Томск»; в Томский государственный университет систем управления и радиоэлектроники.

Таким образом, содержание диссертации последовательное и логичное. Тема и содержание диссертации соответствуют паспорту специальности 05.13.17.

Диссертация является законченным научным трудом и выполнена автором самостоятельно на достаточном научном уровне.

Диссертация содержит достаточное количество исходных данных, имеет пояснения, рисунки, графики, подробные примеры и расчеты. Написана технически квалифицированно и аккуратно оформлена. По каждой главе и работе в целом имеются выводы.

Достоинства и недостатки

Общей положительной стороной диссертации является то, что для решения слабо структурированных проблем автор использует методологию системного анализа, в котором: осуществляет формулировку проблемной ситуации; определяет цель; определяет критерии достижения цели; проводит построение алгоритмов для обоснования решения; проводит поиск допустимых вариантов решения; согласовывает решение; подготавливает решение к реализации; управляет ходом реализации решения; проверяет эффективность решения.

Для решения хорошо структурированных количественно выражаемых проблем автор использует известную методологию исследования операций, которая состоит в построении алгоритмов и применении методов для отыскания оптимальной стратегии управления целенаправленными действиями: проводит абстрагирование, формализацию и конкретизацию; проводит анализ и синтез, использует методы индукции и дедукции; проводит композицию и декомпозицию, структурирование и реструктурирование; осуществляет макетирование; реализует алгоритмизацию; осуществляет моделирование и эксперимент; проводит экспертное оценивание и тестирование.

Замечания по диссертационной работе в целом

1. В первой главе следовало бы больше уделить внимания работам отечественных исследователей в области верификации диктора по произвольной фразе.

2. Отсутствует обоснование применения алгоритма верификации на базе гауссовых смесей и универсальной фоновой модели в качестве базового алгоритма верификации диктора.

3. Отсутствуют рекомендации по выбору фраз, входящих в речевой корпус, используемый для обучения моделей дикторов.

4. Следовало бы подробнее описать практическое использование результатов работы.

Отмеченные недостатки не снижают общей положительной оценки диссертации И.А. Рахманенко и не влияют на корректность полученных автором выводов.

Соответствие автореферата основному содержанию диссертации

Основные этапы научной работы, выводы и результаты представлены в автореферате. Автореферат соответствует основному содержанию диссертации.

Соответствие диссертации и автореферата требованиям ГОСТ Р 7.0.11-2011

В целом оформление диссертации и автореферата соответствует требованиям ГОСТ Р 7.0.11-2011, однако библиографические записи документов в списке литературы диссертации и автореферата оформлены с отступлением от ГОСТ Р 7.0.11-2011.

Заключение

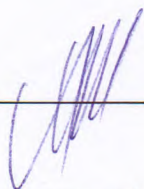
Диссертация Рахманенко Ивана Андреевича является законченной научно-исследовательской работой, выполненной автором самостоятельно на высоком научном уровне и содержащим новые научные результаты. Рассмотренная в диссертационной работе проблема верификации диктора по произвольной фразе соответствует специальности 05.13.17 - «Теоретические основы информатики». В работе приведены научные результаты, позволяющие их квалифицировать как, имеющие существенное значение для развития методов анализа устной речи. Полученные автором результаты достоверны, выводы и заключения обоснованы. Работа базируется на достаточном числе исходных данных, примеров и расчетов. Она написана доходчиво, грамотно и аккуратно оформлена. По каждой главе и работе в целом сделаны четкие выводы.

Основные научные результаты диссертации опубликованы в 10 печатных работах, из которых в рекомендованных ВАК при Минобрнауки РФ периодических изданиях – две. Две работы индексированы в базе научных

публикаций SCOPUS. Автореферат соответствует основному содержанию диссертации.

Диссертационная работа отвечает требованиям п. 9 «Положения присуждения ученых степеней», а ее автор Рахманенко Иван Андреевич заслуживает присуждения ученой степени кандидата наук по специальности 05.13.17 — «Теоретические основы информатики».

Официальный оппонент –
доктор технических наук,
заведующий кафедрой речевых
информационных систем, ФГАОУ ВО
«Санкт-Петербургский национальный
исследовательский университет
информационных технологий,
механики и оптики»



Матвеев Юрий Николаевич

197101, г. Санкт-Петербург, Кронверкский проспект, д.49;
+7 (812) 325-88-48;
matveev@mail.ifmo.ru

Подпись Матвеева Юрия Николаевича удостоверяю:

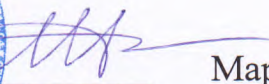
Ученый секретарь Ученого совета ФГАОУ ВО

«Санкт-Петербургский национальный

исследовательский университет

информационных технологий

механики и оптики»



Марусина Мария Яковлевна

30.11.2017г.