

ЗАКЛЮЧЕНИЕ ДИССЕРТАЦИОННОГО СОВЕТА Д 002.073.04

на базе федерального государственного учреждения
«Федеральный исследовательский центр «Информатика и управление»
Российской академии наук» (ФИЦ ИУ РАН) по диссертации
на соискание ученой степени кандидата технических наук

аттестационное дело № _____

решение диссертационного совета от 2 декабря 2019 № 12

О присуждении Булатову Константину Булатовичу, гражданину Российской Федерации, ученой степени кандидата технических наук.

Диссертация «Методы, модели и алгоритмы комбинирования и останова в системах распознавания в видеопотоке» по специальности 05.13.01 «Системный анализ, управление и обработка информации (информационно-вычислительное обеспечение)» принята к защите 12 августа 2019 г., протокол №9, диссертационным советом Д 002.073.04 на базе федерального государственного учреждения «Федеральный исследовательский центр «Информатика и управление» Российской академии наук» (ФИЦ ИУ РАН), 119333, Москва, ул. Вавилова, д. 44, корп. 2, приказ Министерства образования и науки Российской Федерации № 747/нк от 22 июня 2016 г.

Соискатель Булатов Константин Булатович, 1991 года рождения, занимает должность программиста 1 категории ФИЦ ИУ РАН. В 2013 году соискатель получил диплом о высшем образовании с отличием НИТУ «МИСиС», присуждена квалификация «инженер-математик» по специальности «Прикладная математика». С 05.07.2013 по 14.07.2017 обучался в очной аспирантуре НИТУ «МИСиС» и сдал кандидатские экзамены по специальности 05.13.01 – «Системный анализ, управление и обработка информации». С июля 2014 по настоящее время работает в ФИЦ ИУ РАН, с февраля 2018 года состоит в должности программиста 1 категории Отдела 91 «Системное программирование» Отделения 9 «Математическое обеспечение вычислительной техники» ФИЦ ИУ РАН.

Отделения 9 «Математическое обеспечение вычислительной техники»
ФИЦ ИУ РАН.

Диссертация выполнена в Отделе 91 «Системное программирование»
Отделения 9 «Математическое обеспечение вычислительной техники»
ФИЦ ИУ РАН.

Научный руководитель – кандидат технических наук, Арлазаров Владимир
Викторович, заведующий отделом 93 ФИЦ ИУ РАН.

Официальные оппоненты:

1. Мясников Владислав Валерьевич, доктор физико-математических наук, доцент,
Федеральное государственное автономное образовательное учреждение высшего
образования «Самарский национальный исследовательский университет имени
академика С.П. Королева», профессор кафедры геоинформатики и
информационной безопасности.

Отзыв оппонента положительный, имеются замечания по диссертации:

1. Встречаются некорректные или по смыслу неверные формальные
математические выражения. В частности: (а) результат работы системы
распознавания на стр. 39 представлен как отображение из множества
классов в множество вещественных чисел (оценок), что неверно, поскольку
в такой формулировке отсутствует сам объект распознавания. При этом в
последующем выражении (2.1) аргументом выступает уже изображение, а
результатом – множество пар (класс, оценка принадлежности);
(б) формальная запись на странице 40 решения классификатора в виде
$$c^*(I(x)) = \arg \max f^\wedge(I(x))$$
 также некорректна, поскольку результатом
функции $f^\wedge(\dots)$ является указанное выше неупорядоченное множество пар.
2. Встречаются смешения различных понятий и обозначений. В частности:
(а) «Задачу распознавания объекта в видеопотоке можно сформулировать
как поиск классифицирующей функции...» (стр. 46); (б) величины f и $f^\wedge$
одновременно указаны на стр. 56 как «модуль классификации» со
значениями в C (не ясно при этом, в чем их разница), в то время как $f^\wedge$ на

стр. 39 определена как отображение в R ; (в) величина « c » на протяжении разделов 2-3 означала класс, а в разделе 4 – штраф.

3. Предложенные алгоритмы используют числовые параметры (например, параметр α на стр. 49, параметры c , c_e , c_f на стр. 73 и др.). В тексте диссертации не представлены рекомендации по их выбору.

4. Немногочисленные ошибки/описки в тексте: (а) «множеством классом» (стр. 39); (б) «до истинного значения, характеризующая качество» (стр. 44).

2. Нейман-заде Мурад Искендер-оглы, кандидат физико-математических наук, АО «МЦСТ».

Отзыв оппонента положительный, имеются следующие замечания по диссертации:

1. Описанный в главе 2 Функционал эффективности системы содержит параметры a и b , которые выглядят избыточными, т. к. их можно спрятать в нормировку функционала и в штрафную функцию от времени $W(t)$.

2. При исследовании влияния модели входных данных на выбор стратегии комбинирования, наборы тестовых данных были разделены на группы А (с ошибками предварительной обработки) и В (без таковых ошибок), и экспериментально показано, что оптимальные стратегии комбинирования для них отличаются. Этот результат вынесен в список основных результатов диссертационной работы, однако, не проведено и не намечено исследования по автоматической классификации входных данных на группы А и В, с целью построения единого адаптивного механизма выбора стратегии комбинирования.

3. Предложенный в главе 4 новый метод останова существенно опирается на свойство монотонности задачи, при этом не указано, насколько часто (с точки зрения представительных входных данных) это свойство выполняется для задачи последовательного распознавания строк.

4. В предложенном методе останова не рассмотрена возможность обобщения: можно ли предсказывать ожидаемое расстояние не только до следующего неизвестного результата распознавания, но и на несколько шагов вперед.

Ведущая организация – Федеральное государственное учреждение «Федеральный научный центр Научно-исследовательский институт системных исследований Российской академии наук», в своем положительном заключении, подписанном Кушниренко Анатолием Георгиевичем, кандидатом физико-математических наук, заведующим отделом учебной информатики ФГУ ФНЦ НИИСИ РАН, указал, что «Диссертационная работа Булатова Константина Булатовича «Методы, модели и алгоритмы комбинирования и останова в системах распознавания в видеопотоке» является законченной научно-исследовательской работой и соответствует требованиям ВАК РФ, предъявляемым к диссертациям на соискание ученой степени кандидата технических наук по специальности 05.13.01 – Системный анализ, управление и обработка информации (информационно-вычислительное обеспечение), соответствует пункту 3 (Разработка критериев и моделей описания и оценки эффективности решения задач системного анализа, оптимизации, управления, принятия решений и обработки информации) и пункту 4 (Разработка методов и алгоритмов решения задач системного анализа, оптимизации, управления, принятия решений и обработки информации) паспорта специальности. Ее автор, Булатов Константин Булатович, заслуживает присуждения ему ученой степени кандидата технических наук по указанной специальности. Отзыв обсужден и утвержден на семинаре отдела учебной информатики Федерального научного центра Научно-исследовательский институт системных исследований Российской академии наук, протокол № 1 от 7 ноября 2019.»

Замечания:

1. Пункт 6 заключения (стр. 91-92 диссертации) гласит: «6. Результаты работы в качестве программных компонентов систем распознавания документов были внедрены в программное обеспечение «Smart 3D OCR MRZ» и «Smart PassportReader» компании ООО «Смарт Энджинс РУС», а также «Smart IDReader» компании ООО «Смарт Энджинс Сервис». Данные продукты интегрированы в информационную инфраструктуру ряда коммерческих организаций, а также в ряд информационных решений государственных

структур Российской Федерации». Однако упомянутые государственные структуры Российской Федерации остаются неназванными, и в пункте 6 не упоминаются и не названы какие-либо документы, позволяющие подтвердить его справедливость.

Соискатель имеет 9 опубликованных научных работ по теме диссертации, в том числе: 6 статей в рецензируемых научных изданиях, рекомендованных ВАК при Министерстве науки и высшего образования РФ, 3 публикации в сборниках трудов международных научных конференций, входящих в международные базы цитирования Scopus и Web of Science. Соискатель также имеет 2 патента на полезную модель и 3 свидетельства о государственной регистрации программы для ЭВМ. Наиболее значительные работы:

1. Булатов К.Б., Дорохов Е.П., Арлазаров В.В. Алгоритм принятия решения об останове процесса распознавания строки в видеопотоке на основе моделирования следующего интегрированного результата // Информационные процессы, Т. 19, № 2, – 2019, С. 178-186.
2. Арлазаров В.В., Булатов К.Б., Усков А.В. Модель системы распознавания в видеопотоке мобильного устройства // Труды Института системного анализа РАН, Т. 68, Спецвыпуск № S1, – 2018, С. 73-82.
3. Булатов К.Б. Выбор оптимальной стратегии комбинирования покадровых результатов распознавания символа в видеопотоке // Информационные технологии и вычислительные системы, № 3, – 2017, С. 45-55.
4. Булатов К.Б., Кирсанов В.Ю., Арлазаров В.В., Николаев Д.П., Полевой Д.В. Методы интеграции результатов распознавания текстовых полей документов в видеопотоке мобильного устройства // Вестник РФФИ, № 4, – 2016, С. 109-115.
5. Полевой Д.В., Булатов К.Б., Скорюкина Н.С., Чернов Т.С., Арлазаров В.В., Шешкус А.В. Ключевые аспекты распознавания документов с использованием малоразмерных цифровых камер // Вестник РФФИ, № 4, – 2016, С. 97-108.
6. Булатов К.Б., Ильин Д.А., Полевой Д.В., Чернышова Ю.С. Проблемы распознавания машиночитаемых зон с использованием малоформатных цифровых

камер мобильных устройств // Труды Института системного анализа РАН, Т. 65, № 3, – 2015, С. 85-93.

7. Bulatov K., Lynchenko A., Krivtsov V. Optimal frame-by-frame result combination strategy for OCR in video stream // Proc. SPIE (ICMV 2017), – 106961Z, – 2018, DOI: 10.1117/12.2310139

8. Arlazarov V., Bulatov K., Manzhikov T., Slavin O., Janiszewski I. Method of determining the necessary number of observations for video stream documents // Proc. SPIE (ICMV 2017), – 106961X, – 2018, DOI: 10.1117/12.2310132

9. Bulatov K., Arlazarov V., Chernov T., Slavin O., Nikolaev D. Smart IDReader: Document recognition in video stream // Proc. International conference on document analysis and recognition (Camera based document analysis and recognition), V. 6, – 2017, P. 39-44, DOI: 10.1109/ICDAR.2017.347

На автореферат поступило пять положительных отзывов, которые подписали:

1. Миллер Борис Михайлович, доктор физико-математических наук, профессор, главный научный сотрудник Федерального государственного бюджетного учреждения науки Институт проблем передачи информации им. А. А. Харкевича Российской академии наук. Замечания:

1. Недостаточно подробное описание особенностей существующих систем распознавания, а также использование различных терминов «правило останова» и «метод останова» в одном значении (в описании результатов экспериментов из четвертой главы диссертации). Вообще говоря, в теории случайных процессов существует установившийся термин «момент остановки» и теории выбора моментов остановки посвящена многочисленная литература (см. например, работы А. Н. Ширяева, И. Каратзаса и др.)

2. В этих работах для выбора момента остановки используются методы динамического программирования и было бы неплохо сравнить строгие математические подходы с методами, предлагаемыми автором.

2. Чукалина Марина Валерьевна, кандидат физико-математических наук, старший научный сотрудник Института кристаллографии им. А.В. Шубникова ФНИЦ «Кристаллография и фотоника» РАН. Замечаний нет.

3. Свиридюк Георгий Анатольевич, доктор физико-математических наук, профессор, заведующий кафедрой уравнений математической физики ФГАОУ ВО «Южно-Уральский государственный университет (национальный исследовательский университет)». Замечания:

1. Некоторые термины, использующиеся в схеме обработки кадра системы распознавания документа в видеопотоке (рис. 1 автореферата) не определены в тексте автореферата («зона», «пост-обработка»);

2. В автореферате не описано, близка ли задача останова процесса распознавания объекта в видеопотоке к каким-либо из классических постановок задач оптимального останова, и были ли использованы существующие теоретические результаты при построении нового метода, применяющегося к задаче распознавания в видеопотоке.

4. Солодов Сергей Владимирович, кандидат технических наук, и.о. директора института Информационных технологий и АСУ Национального исследовательского технологического университета «МИСиС». Замечания:

1. Подпись к рис. 2, согласно контексту работы, не до конца отражает содержимое рисунка: поскольку не указано, что извлеченные изображения текстового поля (рис. 2, справа), относятся к различным кадрам.

2. В автореферате не указано, останется ли корректным Алгоритм 1, если в качестве функции комбинирования результатов классификации одиночных символов будет использоваться не взвешенное среднее (1), а какая-либо другая стратегия комбинирования.

5. Ногин Владимир Дмитриевич, доктор физико-математических наук, профессор, профессор кафедры теории управления факультета прикладной математики – процессов управления ФГБОУ ВО «Санкт-Петербургский государственный университет». Замечаний нет.

Выбор официальных оппонентов и ведущей организации обосновывается их компетентностью в области обработки цифровых изображений, распознавания образов и построения информационных систем, что подтверждается направлениями их исследований и публикациями в высокорейтинговых научных журналах.

Диссертационный совет отмечает, что на основании выполненных соискателем исследований: разработана новая модель системы распознавания объекта в видеопотоке с блоками комбинирования результатов распознавания одиночных кадров и принятия решения об остановке процесса; **показано** преимущество правила максимальной оценки как стратегии комбинирования покадровых результатов классификации символов в видеопоследовательностях, не содержащих ошибок локализации и сегментации; **разработан** новый алгоритм комбинирования результатов распознавания строк, учитывающий альтернативные варианты классификации отдельных символов; **разработан** новый метод останова процесса распознавания объекта в видеопотоке на основе порогового отсечения оценки ожидаемого расстояния между текущим и следующим интегрированными результатами распознавания; **разработан** новый алгоритм моделирования интегрированного результата распознавания на следующем шаге для применения предложенного метода останова.

Теоретическая значимость исследования обоснована тем, что **введена** новая модель системы распознавания объекта в видеопотоке, **построены** новый алгоритм комбинирования строк, новый метод останова процесса распознавания объекта в видеопоследовательности и алгоритм его применения.

Практическая значимость полученных соискателем результатов исследования заключается в **возможности** практического применения разработанных алгоритмов повышения точности результатов распознавания в видеопоследовательностях в системах распознавания документов. Разработанные методы и алгоритмы комбинирования результатов распознавания строковых объектов и останова процесса распознавания были реализованы в виде программных компонентов и внедрены в программное обеспечение «Smart 3D

