

ОТЗЫВ ОФИЦИАЛЬНОГО ОППОНЕНТА

на диссертационную работу Булатова Константина Булатовича

«Методы, модели и алгоритмы комбинирования и останова в системах распознавания в видеопотоке», представленную на соискание ученой степени кандидата технических наук по специальности 05.13.01 – «Системный анализ, управление и обработка информации (информационно-вычислительное обеспечение)»

Актуальность темы исследований

Развитие технологий и систем анализа цифровых изображений в целом и методов и алгоритмов автоматического распознавания документов в частности являются важными и актуальными задачами, необходимыми для реализации таких государственных программ как Цифровая экономика Российской Федерации (утверждена Правительством РФ в 2017 году) и Национальная стратегия развития искусственного интеллекта (утверждена указом Президента РФ в 2019 году). Диссертационная работа Булатова К.Б. посвящена вопросам разработки новых методов повышения точности результатов распознавания объектов в видеопотоке. Предлагаемые решения применимы к задачам автоматического ввода данных в системах распознавания документов, задачам распознавания объектов в системах видеонаблюдения и др. Таким образом, тема диссертационной работы является актуальной и отвечает современным технологическим вызовам.

Содержание диссертационной работы

Диссертация состоит из введения, четырех разделов, заключения и списка литературы. Общий объем диссертационной работы составляет 109

страниц, включая 18 рисунков и 7 таблиц. Список литературы включает 139 источников.

Во введении приводятся оценка актуальности диссертационной работы, формулировка цели исследования и постановки задач, которые требуется решить для достижения этой цели.

В рамках первого раздела соискателем приводится аналитический обзор существующих принципов и подходов к построению современных систем распознавания документов и задач, возникающих в таких системах. Описываются проблемы, связанные с распознаванием документов на фотографиях и кадрах видеопоследовательности, и приводится обзор методов поиска документа, сегментации изображения документа, распознавания текста, пост-обработки и использования нескольких входных изображений для распознавания. В разделе обоснована актуальность исследований, направленных на создание новых методов и подходов для повышения точности распознавания объектов применительно к системам автоматического ввода документов, оперирующих в реальном времени на мобильных устройствах. В конце раздела формулируются задачи диссертационной работы. Последующие разделы содержат описание полученных соискателем новых научных результатов.

Во втором разделе приводится предлагаемая автором модель системы распознавания объекта в видеопотоке, отличительной особенностью которой является повышение точности результата за счет многократного распознавания изображений объекта на различных кадрах, комбинирования результатов и принятия решения об автоматической остановке этого процесса. В этом же разделе приведены результаты исследования влияния характеристик входных изображений (наличие ошибок локализации и сегментации текста) на выбор оптимального способа комбинирования

результатов классификации изображений символов. Полученные в данном разделе результаты имеют важное значение для построения мобильных систем распознавания и классификации объектов в видеопоследовательностях.

В третьем разделе рассматривается задача комбинирования (объединения) результатов распознавания текстовых строк, полученных в результате анализа отдельных кадров видеопотока. Представлен новый алгоритм комбинирования, который учитывает оценки принадлежности альтернативных результатов посимвольной классификации и продемонстрированы результаты численных экспериментов. Разработанный соискателем алгоритм позволяет увеличить точность распознавания строковых объектов в видеопотоке за счет объединения результатов распознавания индивидуальных кадров.

В четвертом разделе рассматривается задача автоматической остановки процесса покадрового распознавания с комбинированием результатов. Приведена формальная постановка задачи в соответствии с моделью системы, описанной во втором разделе, сформулированы гипотезы о результатах комбинирования результатов и на их основе предложен метод, аппроксимирующий поведение оптимального правила. Для реализации разработанного метода предложен алгоритм моделирования “будущего” комбинирования результата распознавания и представлены результаты численных экспериментов применительно к задаче распознавания текстовой строки. Продемонстрировано значительное преимущество нового метода по сравнению с существующими и ранее опубликованными. Результаты данного раздела имеют важное значение для развития систем распознавания объектов в видеопотоке.

Каждый раздел диссертационной работы содержит четкие выводы, подтверждающие оригинальность и новизну полученных результатов и их практическую значимость.

Степень обоснованности научных положений, выводов и рекомендаций

Все основные положения, выносимые на защиту, являются в достаточной степени обоснованы. Разработанные методы и алгоритмы детально описаны, а достоверность выводов подтверждается численными сравнениями предложенных методов и алгоритмов с известными и ранее опубликованными. Существенная часть экспериментальных исследований производилось с использованием опубликованных пакетов экспериментальных данных. Тем самым представленные экспериментальные результаты являются воспроизводимыми, верифицируемыми и убедительно показывают корректность основных результатов и выводов диссертации. Основные положения и результаты диссертационной работы прошли апробацию на профильных международных конференциях.

Достоверность и научная новизна научных положений и выводов

Научной новизной обладают следующие полученные соискателем результаты:

- модель системы распознавания объекта в видеопотоке, использующая независимое распознавание изображений объектов на отдельных кадрах с последующим комбинированием (объединением) результатов этого распознавания и принятия решения об останове;

- результаты исследования влияния характеристик изображений видеопотока на выбор оптимальной стратегии комбинирования покадровых результатов классификации (в рамках задачи распознавания одиночного символа в видеопотоке);

– алгоритм комбинирования результатов распознавания текстовой строки, учитывающий альтернативные варианты классификации отдельных символов;

– метод останова процесса распознавания объекта в видеопотоке и алгоритм моделирования интегрированного (объединенного) результата распознавания для реализации разработанного метода.

Достоверность положений и выводов подтверждается результатами экспериментальных исследований, представленными в диссертационной работе, а также результатами внедрения разработанных методов и алгоритмов в ряд систем распознавания документов.

Теоретическая и практическая значимость работы

Булатовым К.Б. предложены новая модель системы распознавания объекта в видеопоследовательности, предполагающая комбинирование индивидуальных результатов распознавания изображений объекта на отдельных кадрах с автоматическим принятием решения об остановке процесса; предложен новый алгоритм комбинирования результатов распознавания строковых объектов и разработан новый метод останова процесса распознавания объекта в видеопотоке. Полученные результаты обладают как теоретической, так и практической значимостью для построения систем распознавания в видеопотоке, в частности, систем распознавания и автоматического ввода документов, оперирующих в реальном времени на мобильных устройствах.

Представленные в диссертационной работе методы и алгоритмы реализованы соискателем в виде программных компонентов и внедрены в программные комплексы распознавания документов, удостоверяющих личность, Smart 3D OCR MRZ, Smart PassportReader и Smart IDReader.

Публикация основных результатов диссертации

Основное содержание диссертационной работы представлено в 14 печатных публикациях, из которых 9 представляют собой публикации в журналах и сборниках трудов профильных международных конференций. При этом 6 работ опубликовано в журналах, входящих в рекомендованный ВАК список, а 3 работы представлены в трудах международных конференций, индексируемых Scopus и/или Web of Science.

Опубликованные работы и автореферат полностью отражают содержание диссертации.

Замечания

Диссертационная работа Булатова К.Б. в полной мере отражает решение поставленных задач и подтверждает достижение сформулированной цели. По тексту работы имеются следующие замечания.

1. Встречаются некорректные или по смыслу неверные формальные математические выражения. В частности:

- 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);

- формальная запись на странице 40 решения классификатора в виде

$$c^*(I(x)) = \arg \max \hat{f}(I(x))$$

также некорректна, поскольку результатом функции $\hat{f}(\dots)$ является указанное выше неупорядоченное множество пар.

2. Встречаются смешения различных понятий и обозначений. В частности:

- "Задачу распознавания объекта в видеопотоке можно сформулировать как поиск классифицирующей функции..." (стр. 46);

- величины f и $\hat{f}$ одновременно указаны на стр. 56 как "модуль классификации" со значениями в C (не ясно при этом, в чем их разница), в то время как $\hat{f}$ на стр.39 определена как отображение в $\mathbf{R}$;

- 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:

- "множеством классом..." (стр. 39);

- "до истинного значения, характеризующая качество" (стр. 44).

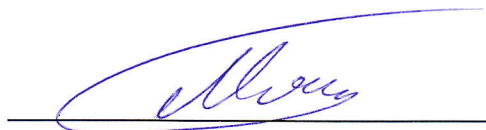
Как правило, указанные ошибки для читателя устранимы по содержательному контексту работы.

Общая оценка работы

Перечисленные недостатки и замечания не снижают общую положительную оценку диссертационной работы Булатова К.Б., ее научную ценность и практическую значимость. Диссертационная работа Булатова К.Б. является законченной научно-квалификационной работой, отвечающей критериям Положения о присуждении ученых степеней, утвержденного постановлением Правительства Российской Федерации от 24.09.2013 г. № 842 (с изменениями постановлений Правительства РФ от 30.07.2014 г. № 723, от 21.04.2016 г. № 335, от 02.08.2016 г. № 748, от 29.05.2017 г. № 650, от 28.08.2017 г. № 1024, от 01.10.2018 г. № 1168), предъявляемым к кандидатским диссертациям на соискание ученых степеней.

Считаю, что Булатов Константин Булатович заслуживает присуждения ученой степени кандидата технических наук по специальности 05.13.01 – «Системный анализ, управление и обработка информации (информационно-вычислительное обеспечение)».

профессор кафедры геоинформатики и информационной безопасности федерального государственного автономного образовательного учреждения высшего образования «Самарский национальный исследовательский университет имени академика С.П. Королева», д.ф.-м.н., доцент



/ В.В. Мясников

«06» ноября 2019 г.

Мясников Владислав Валерьевич,
федеральное государственное автономное образовательное учреждение высшего образования «Самарский национальный исследовательский университет имени академика С.П. Королева», кафедра геоинформатики и информационной безопасности,
443086, г. Самара, ул. Московское шоссе, д. 34,
+7 (846) 267-43-70, vmyas@geosamara.ru

