

ОТЗЫВ ОФИЦИАЛЬНОГО ОППОНЕНТА НА ДИССЕРТАЦИЮ

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 диссертационной работы

Для решения задач автоматизации в последние годы растет доля используемых малоформатных цифровых камер и компактных вычислительных устройств, таких как смартфоны и планшетные компьютеры. Системы обработки, анализа и распознавания изображений для такого класса устройств имеют важное значение. Изображения, полученные при помощи малоформатных цифровых камер, обладают рядом характерных проблем, таких как недостаточный контраст, недостаточный уровень сфокусированности на объекте, и другие. При этом в реальных задачах, при видеосъемке объекта, существует возможность обработки нескольких кадров и тем самым повысить точность извлечения информации. Основная часть диссертационной работы посвящена методам повышения точности распознавания при использовании нескольких образов одного и того же объекта (задача комбинирования) и методам оптимальной остановки процесса распознавания объекта на видео.

Структура работы и ее содержание

Диссертационная работа соискателя состоит из введения, четырех глав, заключения, списка литературы, списка рисунков и списка таблиц. Общий объем

диссертации составляет 109 страниц, список литературы состоит из 139 источников.

Во введении диссертационной работы приведено обоснование актуальности темы работы, формулируется цель диссертационного исследования и приводятся поставленные задачи, решение которых необходимо для достижения поставленной цели. Приведена достаточная аргументация научной новизны проведенного исследования, показывается практическая значимость полученных результатов.

В первой главе работы соискатель приводит анализ принципов построения современных систем распознавания, фокусируясь на системы распознавания документов как на важнейшую составляющую систем автоматизированного документооборота. В главе показано, что современные системы автоматического ввода и распознавания документов с помощью мобильных устройств рассматривают лишь фотографию документа (единственное изображение) и отмечают сложности с предварительной обработкой изображения и с самим распознаванием. В главе также приведены результаты существующих исследования методов комбинирования результатов распознавания одного изображения несколькими системами распознавания и указано, что подходам к комбинированию результатов распознавания различных изображений одного и того же объекта в литературе уделяется недостаточно внимания. Также в главе приведены результаты исследования методов решения задачи останова и описана так называемая “задача о вычитке”, которую по ее постановке можно считать наиболее близкой к задаче останова процесса распознавания объекта в видеопотоке, которой будет посвящена четвертая глава диссертационной работы. Таким образом, с учетом приведенного обзора важность и оригинальность предлагаемого исследования в работе хорошо обоснована.

Во второй главе Булатов К.Б. предлагает оригинальную модель системы оптического распознавания объекта в видеопотоке с блоком комбинирования

результатов, полученных для одиночных изображений, и с блоком останова процесса распознавания. Для построенной модели приведена математическая постановка задачи оптимизации функционала эффективности такой системы как линейной комбинации ошибки распознавания (в терминах расстояния от комбинированного результата до истинного значения) и времени, необходимого для получения этого результата. Рассмотрена задача комбинирования покадровых результатов распознавания объектов, показаны ее сходства с задачей “ансамбля” – комбинирования результатов распознавания одного и того же изображения несколькими методами. Приведены результаты экспериментального исследования, направленного на выбор наиболее эффективной стратегии комбинирования результатов классификации символа в условиях наличия или отсутствия шумовых помех во входных данных. Также поставлена новая задача останова процесса распознавания видеопотока и обозначена ее важность в контексте систем распознавания, работающих в реальном времени на мобильных устройствах.

В третьей главе автор решает задачу комбинирования нескольких результатов распознавания текстовой строки в видеопотоке. Описан новый метод, основанный на подходе ROVER (который ранее использовался в задаче ансамбля систем распознавания речи), модифицированный для использования в модели результата распознавания строки как последовательности результатов классификации отдельных символов. В главе подробно описаны все этапы построения алгоритма и его основные свойства. Соискатель приводит результаты экспериментального исследования предложенного алгоритма для решения задачи комбинирования результатов распознавания текстовых строк на идентификационных документах и показывает, что предложенный алгоритм превосходит подход ROVER (который не использует альтернативные варианты классификации отдельных символов) по достигнутой средней точности комбинированного результата. Таким образом, основным результатом

исследований, описанных в главе, является новый алгоритм комбинирования результатов распознавания строки в видеопотоке, позволяющий значительно повысить точность результата распознавания в видеопотоке.

В четвертой главе Булатов К.Б. рассматривает новую задачу останова процесса распознавания объекта в видеопотоке. Приводится формальная постановка задачи, определение понятия “монотонных задач останова” и формулируется требование к алгоритму комбинирования результатов распознавания, при выполнении которого задача останова становится монотонной начиная с некоторого шага. После этого предлагается оригинальный метод останова, использующий оценку ожидаемой разницы ошибок на текущем и следующем шагах при помощи неравенства треугольника. Для реализации метода предлагается алгоритм моделирования “следующего” комбинированного результата. Приведены результаты экспериментального исследования и показано, что при помощи предложенного алгоритма удастся достигнуть более высокой эффективности, чем тривиальные и ранее опубликованные методы останова (т.е. достигнуть меньшего уровня средней ошибки при том же среднем количестве обработанных кадров и меньшее среднее количество обработанных кадров при том же уровне средней ошибки).

В заключении приводятся основные результаты диссертационной работы, в том числе сводный результат работы системы распознавания, в которой реализованы как алгоритм комбинирования покадровых результатов, описанный в третьей главе, так и метод останова процесса распознавания в видеопотоке, описанный в четвертой главе.

Научная новизна работы

Основная новизна диссертационной работы состоит в следующем:

1. Предложена новая модель системы распознавания объекта в видеопотоке, состоящая из модулей распознавания одиночных изображений, комбинирования покадровых результатов и оценки критерия останова.

2. Разработан и описан новый алгоритм комбинирования результатов распознавания строковых объектов (таких, как текстовые строки), учитывающий альтернативные варианты классификации отдельных компонентов (символов).

3. Разработан новый метод останова процесса распознавания объекта в видеопотоке, рассматривающий данный процесс как монотонную задачу останова и применяющий аппроксимацию близорукого правила.

Степень обоснованности и достоверности результатов исследования

Все положения диссертации детально описаны и научно обоснованы. Все выводы напрямую следуют из описанных положений. Достоверность результатов, полученных соискателем, подтверждается экспериментальными данными, представленных в диссертационной работе.

Практическая значимость

Разработанные Булатовым К.Б. методы и алгоритмы реализованы в виде программных компонентов и используются в рамках программного обеспечения Smart 3D OCR MRZ, Smart Passport Reader, и Smart IDReader компании ООО «Смарт Энджинс Сервис». В автореферате диссертации приведены данные свидетельств о государственной регистрации данного программного обеспечения, соискатель указан в качестве соавтора. Методы и алгоритмы комбинирования и останова, разработанные в рамках диссертации, могут применяться как в других системах распознавания, использующих множественные изображения, так и в других системах компьютерного зрения при должном расширении или проведении дополнительных экспериментальных исследований.

Рекомендации по использованию результатов работы

Полученные в рамках диссертационной работы результаты представляют большой теоретический и практический интерес, и могут быть использованы для дальнейшего изучения и развития науки в этой области. Экспериментальные методики, использованные автором в рамках работы, могут быть использованы для дальнейшего развития методов комбинирования результатов распознавания и исследования методов останова процесса распознавания объектов различной природы и в различных условиях.

Публикация основных результатов диссертации

Основные результаты диссертационной работы представлены в 14 печатных публикациях, из которых 6 являются статьями, изданными в журналах, рекомендованных ВАК, 3 – в сборниках трудов конференций, входящих в международные индексы цитирования Scopus и Web of Science, 2 – патенты на полезную модель и 3 – свидетельства о государственной регистрации программы для ЭВМ. Опубликованные работы и автореферат полностью отражают содержание диссертации и выполнены на высоком научном уровне.

Замечания

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.

Указанные замечания не являются критическими, поскольку отмеченные недостатки не нарушают целостность работы и значительным образом не сказываются на корректности полученных результатов.

Заключение

На основании представленных материалов диссертации считаю, что диссертация Булатова Константина Булатовича «Методы, модели и алгоритмы комбинирования и останова в системах распознавания в видеопотоке» является законченным научно-квалификационным трудом и удовлетворяет требованиям Положения о порядке присуждения ученых степеней, а ее автор, К.Б. Булатов, заслуживает присуждения искомой ученой степени кандидата технических наук по

специальности 05.13.01 – «Системный анализ, управление и обработка информации (информационно-вычислительное обеспечение)».

Официальный оппонент:

Начальник отделения “Языки программирования,
оптимизирующие компиляторы на различные
платформы” АО «МЦСТ»

кандидат физико-математических наук
(специальность 01.01.01- Вещественный,
комплексный и функциональный анализ).

Нейман-заде М.И.



Сведения об организации:

АО «МЦСТ»

Адрес: 117105, Москва, ул. Нагатинская, д. 1, стр.23

тел: +7 (495) 363 96 65

факс: +7 (495) 363 95 99

E-mail: mcst@mcst.ru

Подпись кандидата физико-математических наук

Нейман-заде М.И. заверяю

Зам. генерального директора АО «МЦСТ» по науке,

доктор технических наук, профессор



Перекатов В.И.

14.11.2019