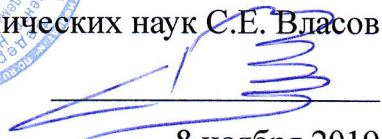




«УТВЕРЖДАЮ»

Директор ФГУ ФНЦ НИИСИ РАН
Доктор технических наук С.Е. Власов


8 ноября 2019

ОТЗЫВ

ведущей организации 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)

Работа с видеопотоком в последние годы является одной из наиболее актуальных и активно развивающихся областей компьютерного зрения. Среди задач интеллектуальной обработки видео интерес вызывают такие задачи, как детектирование объектов на кадрах и слежение за объектами в реальном времени, анализ изменения образов объектов с течением времени, распознавание в видеопотоке объектов различной природы, как ригидных, так и изменяющих свою форму. Системы интеллектуальной обработки видео находят свое применение в робототехнике, мониторинге технологических, производственных или естественных природных процессов, и в других областях. Развитие персональных мобильных устройств, таких как мобильные телефоны с интегрированными цифровыми камерами, привело к новому витку развития систем компьютерного зрения, обрабатывающих фотографии, полученные при помощи мобильных

устройств, и способных, наряду с обработкой отдельных кадров, решать задачи детектирования, анализа и распознавания объектов в видеопотоке в реальном времени.

Диссертация Булатова Константина Булатовича «Методы, модели и алгоритмы комбинирования и останова в системах распознавания в видеопотоке» направлена на развитие систем распознавания объектов в видеопотоке. В рамках диссертации проведено исследование задач компьютерного зрения, связанных с распознаванием документов с применением мобильных устройств и выявлен комплекс нерешенных задач, а именно: задача оптимального комбинирования результатов распознавания, происходящего в реальном времени с целью повысить точность за счет многократного распознавания различных образов одного и того же объекта, и задача оптимального останова процесса распознавания объекта в видеопотоке. Предложена новая модель, а также набор методов и алгоритмов, позволяющих решать такие задачи, проведен ряд вычислительных экспериментов, показывающий их эффективность. В ходе диссертационного исследования Булатовым К.Б. были получены научные и практические результаты, значимые как для дальнейшего исследования систем распознавания образов, так и для практического применения таких систем для решения производственных, технических и социальных задач.

К основным результатам диссертационной работы Булатова Константина Булатовича можно отнести следующее:

1. Предложена новая модель системы распознавания объекта в видеопоследовательности, предполагающая независимое распознавание объекта на отдельных кадрах, комбинирование независимо полученных результатов и принятие решения об остановке данного процесса.
2. Проведено исследование влияние шума входных данных на выбор оптимальной стратегии комбинирования результатов классификации одиночных объектов (таких, как печатные символы).

3. Разработан новый алгоритм комбинирования результатов распознавания строки для случая посимвольной классификации, учитывающий возможные неточности сегментации изображений строк на отдельные символы (и, следовательно, необходимость комбинирования строк разной длины).
4. Разработан новый метод решения задачи оптимального останова процесса распознавания объекта в видеопоследовательности, основанный на оценке ожидаемого расстояния между текущим комбинированным результатом и комбинированным результатом на следующем шаге.
5. Предложен алгоритм моделирования следующего комбинированного результата распознавания объекта в видеопоследовательности, необходимый для реализации разработанного метода решения задачи останова.

Общая характеристика диссертационной работы. Диссертация имеет общий объем 109 страниц, состоит из введения, четырех глав, заключения, списков литературы, рисунков и таблицы. Список литературы содержит 139 наименований.

Во введении обосновывается актуальность выбранной темы,

В первой главе диссертации автор приводит подробный анализ принципов построения и работы современных систем распознавания документов. Выделены и описаны основные этапы обработки изображений документов, как полученных путем сканирования, так и полученных при помощи малоформатных цифровых камер. Показано, что большинство современных работ, посвященных автоматическому вводу, анализу и распознаванию документов, рассматривает одиночное изображение (фотографию или скан) документа как его электронное представление, а также отмечают сложности, связанные с подготовкой образа документа к распознаванию и с самим распознаванием. В качестве вывода ставится задача рассмотрения видеопотока как цифрового представления документа и

описываются возникающие в этой связи задачи комбинирования результатов распознавания.

Во второй главе предлагается новая модель системы распознавания объекта в видеопотоке. Отличительными чертами модели является наличие компонентов, отвечающих за комбинирование результатов распознавания объекта на одиночных кадрах видеопотока (блок интеграции, модуль памяти), и компонента принятия решения об останове – т.е. о том, что процесс захвата новых кадров можно прекратить и накопленный к текущему моменту результат можно принять за окончательный. Описывается функционал эффективности системы распознавания объекта с модулями комбинирования и останова, как линейная комбинация точности комбинированного результата (в терминах расстояния до истинного значения) и временем, затраченным на получение этого результата (в рамках диссертационной работы под «временем» понимается просто количество обработанных кадров).

Во второй главе приводятся также результаты исследования, отвечающего на вопрос о том, каким образом наличие или отсутствие ошибок предварительной обработки входных изображений влияет на выбор оптимальной стратегии комбинирования результатов распознавания. В качестве модельной системы для проведения исследования автором рассматривалась система классификации одиночного символа, на вход которой подается последовательность изображений одного и того же символа, либо полученные точной локализацией объекта на разных кадрах видеопотока, либо с привнесенными на некоторых кадрах ошибками локализации. Продемонстрирована значительная разница в оптимальном выборе стратегии комбинирования в зависимости от модели входных данных: в экспериментах с наличием ошибок локализации объекта более высокую точность обеспечивало правило произведения максимальных оценок классификации, тогда как в экспериментах с точной локализацией на всех кадрах более высокую точность обеспечивало правило максимальной оценки. Результат данного исследования

имеет важное значение для применения на практике модели распознавания объекта в видеопотоке, предложенной автором в данной главе.

Третья глава диссертационной работы посвящена проблеме комбинирования результатов распознавания строковых объектов — т.е. цепочек результатов классификации отдельных компонентов. Помимо необходимости выбора стратегии комбинирования результатов классификации, в такой задаче также необходимо решать задачу сопоставления компонентов результатов, полученных обработкой разных кадров. В случае, когда результатом классификации каждого отдельного компонента является либо одиночная метка класса, либо пара (метка класса, оценка), данная задача уже рассматривалась в литературе. Однако более общий случай, в котором результатом классификации одиночного объекта является множество оценок, соответствующих каждому допустимому классу, ранее не рассматривался. В главе описан новый алгоритм комбинирования результатов распознавания строковых объектов, который применим к модели результата распознавания с альтернативными результатами классификации компонентов. Приведены результаты экспериментального исследования, целью которого было сравнение результатов работы нового алгоритма (в терминах достигаемой точности результата) с результатами алгоритмов, основанных на ранее опубликованных методах, использующих лишь метку класса с максимальной оценкой. Автором убедительно продемонстрировано преимущество разработанного алгоритма перед существующими аналогами.

В четвертой главе рассматривается задача принятия решения об остановке процесса комбинирования покадровых результатов. Задача оптимального останова является одной из классических задач математической статистики и теории принятия решения, однако применительно к системам распознавания она является новой и малоизученной. Процесс распознавания объекта в видеопотоке трактуется в этой главе как процесс последовательного принятия решений. Приводится формальная постановка задачи оптимального останова и предлагается метод ее

решения в рамках такой модели, не зависящий от выбранного алгоритма комбинирования индивидуальных покадровых результатов и от функции близости комбинированного результата и истинного значения (при условии, что функция близости удовлетворяет неравенству треугольника). Разработанный метод заключается в оценке математического ожидания расстояния от комбинированного результата распознавания на текущем шаге до такового на следующем шаге, с последующим пороговым отсечением этой оценки. Для реализации метода предлагается алгоритм моделирования следующего комбинированного результата, исходящий из предположения о том, что новое наблюдение будет близко к уже полученным на предыдущих этапах результатам покадрового распознавания. Далее в главе приводятся результаты экспериментального исследования, показывающие преимущество нового метода и алгоритма перед ранее существующими и опубликованными аналогами: используя предложенный автором диссертации метод при одном и том же среднем количестве наблюдений можно достигнуть большей средней точности результата распознавания в момент останова, а при зафиксированной средней точности распознавания – меньшего в среднем количества обрабатываемых кадров.

В заключении диссертации приводятся основные результаты диссертационной работы.

Достоверность научных результатов обеспечивается результатами проведенных соискателем вычислительных экспериментов.

Диссертация написана достаточно хорошим языком, содержит подробный анализ существующих работ и богатую библиографию. Основные результаты автора изложены ясно. В период с 2017 по 2019 гг. основные результаты диссертационной работы докладывались на двух международных конференциях: на секции по анализу и распознаванию документов на изображениях, полученных с камеры Международной конференции по анализу и распознаванию документов (2017 г., Киото, Япония), и на Международной конференции по машинному зрению

(2017 г., Вена, Австрия), а также на Международном научно-исследовательском семинаре «Анализ и понимание изображений (Математические, когнитивные и прикладные проблемы анализа изображений и сигналов)» (2019 г., Москва, РФ).

Основные результаты диссертационной работы опубликованы в открытой печати: 6 статей в журналах, рекомендованных ВАК, 3 публикации в сборниках трудов конференций, входящих в международные базы цитирования Scopus и Web of Science, 3 свидетельства на программу для ЭВМ и 2 патента на полезную модель.

В диссертации может быть отмечен следующий недостаток:

Пункт 6 заключения (стр. 91-92 диссертации) гласит:

«6. Результаты работы в качестве программных компонентов систем распознавания документов в видеопотоке были внедрены в программное обеспечение «Smart 3D OCR MRZ» и «Smart PassportReader» компании ООО «Смарт 92 Энджинс РУС», а также «Smart IDReader» компании ООО «Смарт Эндж инс 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.

Указанное замечание по диссертации не влияет на положительную оценку диссертации в целом и не ставят под сомнение полученные автором результаты. Результаты, полученные в результате проделанной соискателем работы, своевременно опубликованы и представляют интерес как в научном, так и в практическом плане. Автореферат диссертации полностью отражает ее содержание.

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.

Зав. отделом учебной информатики
ФГУ ФНЦ НИИСИ РАН,
кандидат физико-математических наук

_____ *Куз* А.Г. Кушнirenко
« *8* » *ноября* 2019 г.

Адрес ведущей организации:
17218, Москва, Нахимовский просп., 36, к.1.
Федеральное государственное учреждение
"Федеральный научный центр
Научно-исследовательский институт
системных исследований
Российской академии наук"
e-mail: niisi@niisi.msk.ru



Подпись руки *А.Г. Кушнirenко* заверяю
Начальник отдела кадров *Юлия Полякова*