

ОТЗЫВ

на автореферат диссертации Смирнова Сергея Владимировича «Технология и система автоматической корректировки результатов при распознавании архивных документов», представленной на соискание ученой степени кандидата технических наук по специальности 05.13.11 — «Математическое и программное обеспечение вычислительных машин, комплексов и компьютерных сетей»

В последние годы активно набирают темпы процессы оцифровки документов культурного и исторического наследия, большой процент таких документов содержит машинописный текст, успешно подвергающийся оптическому распознаванию. Полученные результаты распознавания далее передаются в системы полнотекстовой индексации и используются для оперативного поиска и навигации по созданной базе данных. Наглядным примером использования результатов оцифровки исторических документов является сформированный информационный ресурс: электронный банк документов «ПОДВИГ НАРОДА В ВЕЛИКОЙ ОТЕЧЕСТВЕННОЙ ВОЙНЕ 1941-1945 ГГ.» Очевидно, что ошибочная идентификация в содержании исходного документа, допущенная в процессе распознавания, неизбежно приведет к ошибкам в поисковых результатах и существенно снизит возможность обнаружения требуемых документов исследователями.

По указанным причинам, диссертационная работа Смирнова С.В., направленная на решение задач разработки технологии и программно-технических решений, обеспечивающих повышение качества распознавания архивных документов и корректировки допущенных ошибок является актуальной и представляет несомненный практический интерес.

Автором, на наш взгляд, правильно определены объект и предмет исследования.

Среди наиболее значимых научных результатов, полученных автором в ходе диссертационных исследований целесообразно выделить следующие:

- предложен метод автоматической корректировки результатов оптического распознавания документов различных тематических областей, адаптированный под специфику архивной отрасли и не требующий участия человека на стадиях обучения системы и выборки наилучших корректировок;
- разработана система для проведения массового потокового распознавания большого количества электронных образов архивных документов с автоматической корректировкой результатов, предоставляющая возможность детальной настройки всех процессов обработки и оценки их качества при помощи специального инструментария;
- предложена новая технология распознавания крупных массивов документов, содержащая четкие методические инструкции и регламентирующая весь процесс обработки всего корпуса изображений.

Теоретическая значимость полученных научных результатов заключается в развитии методов автоматической корректировки результатов оптического распознавания документов различных тематических областей.

Практическая ценность диссертационных исследований заключается в том, что реализация полученных научных результатов позволила сформировать новую технологию распознавания крупных массивов документов, ориентированную на специфические особенности архивного документооборота.

Результаты проведенных автором исследований и разработок нашли практическое применение и были апробированы в пяти центральных государственных архивах и Архивном комитете Санкт-Петербурга.

Основные положения и результаты диссертационной работы прошли достаточную апробацию на научно-технических конференциях различных уровней и отражены в 14 публикациях, из которых 6 статей в рецензируемых научных изданиях из перечня ВАК, 2

свидетельства о государственной регистрации программ для ЭВМ, что выгодно отличает работу.

Исходя из содержания автореферата, можно считать, что поставленная цель – повышение качества распознавания архивных документов, автором достигнута.

Несмотря на общее положительное впечатление от данной работы имеется ряд замечаний:

1. В автореферате не отмечены пути дальнейшего развития полученных научных результатов и возможности их совместного использования с достижениями других областей науки.
2. При описании предложенной технологии (стр. 13) имеются расхождения в представлении этапов на рисунке 3 и описанием этих этапов по тексту.
3. На странице 18 приведен рисунок 9 «Процентное распределение количества изображений по диапазонам словарной точности результатов распознавания до и после корректировки». К рисунку нет детальных пояснений, что затрудняет его понимание и необоснованно лишает наглядности высокую эффективность произведенной корректировки.

Отмеченные недостатки не носят принципиального характера и не снижают научную и практическую значимость полученных результатов.

Судя по автореферату, диссертационная работа Смирнова Сергея Владимировича является законченной научно-квалификационной работой, удовлетворяет требованиям «Положения о порядке присуждения ученых степеней», предъявляемым ВАК РФ к кандидатским диссертациям по специальности 05.13.11 — «Математическое и программное обеспечение вычислительных машин, комплексов и компьютерных сетей».

Соискатель Смирнов Сергей Владимирович заслуживает присуждения ученой степени кандидата технических наук.

Доцент кафедры защиты информации и систем связи
Федерального государственного бюджетного образовательного учреждения высшего профессионального образования "Юго-Западный государственный университет",
кандидат технических наук,
старший научный сотрудник

Почтовый адрес: г. Курск, ул. 50 лет Октября, д. 94

Телефон: (4712) 54 59 73


Электронная почта: tk_kursk@mail.ru

 Сергей Николаевич Михайлов


«21» апреля 2015 г.

Отзыв рассмотрен и одобрен на заседании кафедры ЗИ и СС, протокол №12 от 22.04.2015 г.

Заведующий кафедрой «ЗИ и СС»
кандидат технических наук,
старший научный сотрудник

 Александр Михайлович Потапенко



 Алексеева С.М.
Подпись
Удостоверяю
Специалист в/о кадрам

 Потапенко А.М.

 Потапенко А.М.

 Потапенко А.М.