

## 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 настоящее время активно развивается сфера предоставления государственных электронных услуг, эффективное оказание которых требует автоматизации деятельности многих ведомств. Применительно к архивной отрасли одной лишь автоматизации рабочих процессов оказывается не достаточно, поскольку основной задачей в данном случае является поиск информации. На первый план выходят задачи перевода бумажных фондов в электронный вид и последующее оптическое распознавание. Однако существующие OCR системы, в большинстве своем, направлены на обработку современных документов, и их точность заметно падает при распознавании исторических архивных материалов, качество которых может быть значительно хуже.

В связи с этим работа Смирнова С.В., посвященная разработке научно-обоснованных программно-технических решений для повышения качества процессов оптического распознавания архивных документов путем применения метода, технологии и системы автоматической корректировки результатов, является, несомненно, актуальной и обладает высокой практической значимостью.

Автором диссертационной работы был получен ряд новых научных результатов, наиболее интересные из которых следующие:

Метод автоматической корректировки результатов оптического распознавания, основанный на рейтинго-ранговой модели текста, формируемой путем лингвостатистического анализа всего корпуса распознанных материалов. Важным конкурентным преимуществом данного метода является способность обрабатывать большие массивы документов разнообразных тематических областей без предварительного обучения и участия человека на этапе выбора корректировок.

Технология и система распознавания архивных документов с последующей корректировкой результатов посредством разработанного метода, предназначенные для массовой потоковой обработки больших массивов электронных образов и автоматизирующие весь процесс от предобработки изображений до многокритериальной оценки качества полученных результатов.

Инструментарий настройки конфигурационных профилей системы для обработки различных групп документов, отличающихся качеством, тематикой или другими характеристиками.

Достоинством работы Смирнова С.В. является то, что все полученные результаты были испытаны на практике при обработке документов пяти центральных государственных архивов Санкт-Петербурга в составе государственной информационной системы «Государственные архивы Санкт-Петербурга». Проведенная опытная эксплуатация наглядно продемонстрировала эффективность предложенных автором решений.

К автореферату имеются следующие замечания:

1. Наглядность формализованного описания метода автоматической корректировки можно было бы улучшить путем иллюстрирования реальными примерами.
2. Также хотелось бы увидеть пример графического интерфейса разработанного программного модуля сравнительного анализа.

Несмотря на отмеченные недостатки, рассмотренная работа, судя по автореферату, представляет завершённый научно-квалификационный труд, в котором решена важная и актуальная задача повышения эффективности оптического распознавания архивных документов.

Диссертация отвечает требованиям Положения ВАК РФ о порядке присуждения ученых степеней, предъявляемым к кандидатским диссертациям, а Смирнов Сергей Владимирович заслуживает присуждения ученой степени кандидата технических наук по специальности 05.13.11 — «Математическое и программное обеспечение вычислительных машин, комплексов и компьютерных сетей».

Почетный профессор Федерального государственного образовательного бюджетного учреждения высшего профессионального образования «Санкт-Петербургский государственный университет телекоммуникаций им. проф. М.А. Бонч-Бруевича»,  
 профессор кафедры «Информатики и компьютерного дизайна»,  
 доктор технических наук, профессор  
 Дегтярев Владимир Михайлович  
 18 мая 2015 г.

Почтовый адрес: 191186, Санкт-Петербург, наб. реки Мойки, д. 61, СПбГУТ им. проф. М.А. Бонч-Бруевича  
 Телефон: (812) 3235176  
 Электронная почта: [dvm1940@gmail.com](mailto:dvm1940@gmail.com)

