

**МИНИСТЕРСТВО ОБРАЗОВАНИЯ И НАУКИ
РОССИЙСКОЙ ФЕДЕРАЦИИ**

федеральное государственное бюджетное образовательное
учреждение высшего образования

**«Сибирский государственный
университет науки и технологий
имени академика М.Ф. Решетнева»
(СибГУ им. М.Ф. Решетнева)**

проспект им. газеты Красноярский рабочий, 31
г. Красноярск, 660037
тел.: +7 (391) 264-00-14 факс: +7 (391) 264-47-09
<http://www.sibsau.ru> e-mail: info@sibsau.ru
ОКПО 02069734, ОГРН 1022402056038
ИНН/КПП 2462003320/246201001

08 ДЕК 2017 № 73/5-5633

На № _____ от _____

«УТВЕРЖДАЮ»

Проректор по научной и
инновационной деятельности,
д-р физ.-мат. наук

Ю.Ю. Логинов

« 8 декабря 2017 г.



ОТЗЫВ ВЕДУЩЕЙ ОРГАНИЗАЦИИ

**Федерального государственного бюджетного образовательного
учреждения высшего образования «Сибирский
государственный университет науки и технологий имени
академика М.Ф. Решетнева»**

на диссертацию Хаустова Павла Александровича «Алгоритмы распознавания рукописных символов в условиях малой обучающей выборки», представленную на соискание учёной степени кандидата технических наук по специальности 05.13.11 – Математическое и программное обеспечение вычислительных машин, комплексов и компьютерных сетей

Актуальность темы диссертационной работы

Тема работы относится к области распознавания рукописных текстов в условиях малого количества эталонных изображений. Задача распознавания рукописных текстов активно исследуется в течение многих лет, современные методы её решения преимущественно основаны на применении систем машинного обучения, в частности, свёрточных нейронных сетей. Такие методы требуют наличия объёмных обучающих выборок, что затрудняет их применение в случаях, когда подготовить такие выборки не представляется возможным.

Сложность рассматриваемой задачи обусловлена большим многообразием почерков и форм начертаний символов, произвольным исполнением отдельных элементов графического представления, различными способами соединения соседних символов в слове. При этом далеко не всегда малочисленный набор эталонных изображений символов позволяет задать всё многообразие форм представления символов и допустимых видоизменений отдельных элементов начертания.

Целью диссертационного исследования Хаустова П.А. является разработка алгоритмов распознавания рукописных символов, которые

способны функционировать в условиях, когда крайне затруднительно найти и получить большое количество эталонных изображений для обучения или настройки классификатора, что является актуальной задачей.

Структура работы и её содержание

На отзыв ведущей организации была представлена диссертационная работа, состоящая из введения, четырёх глав, заключения, списка литературы из 144 источников и двух приложений. Текст диссертации изложен на 181 странице, сопровождается 80 рисунками и 30 таблицами.

Во *введении* обосновывается актуальность темы исследования, формулируются цель и основные задачи диссертационной работы. Описываются теоретическая и практическая значимость полученных результатов. Обоснована научная новизна исследования и представлены основные положения, выносимые на защиту.

В *первой* главе представлен аналитический обзор ранее предложенных методов и алгоритмов, применяемых для распознавания символов и обсуждаются однопроходные алгоритмы скелетизации бинарного представления символов.

Во *второй* главе автор описывает предлагаемый им алгоритм распознавания рукописных символов в условиях малой обучающей выборки на основе построения структурных моделей. Алгоритм состоит из трёх стадий. Первая стадия заключается в предварительной обработке с целью устранения дефектов сканирования, бинаризации и последующей скелетизации изображения. Предложенный алгоритм скелетизации бинарных изображений символов заключается в последовательном применении двух общеизвестных однопроходных алгоритмов скелетизации: Зонга-Суня и Ву-Цая. Для устранения элементов начертания, которые приводят к неправильной работе алгоритма Зонга-Суня, предложено использовать операцию предварительной однопроходной обработки. Вторая стадия алгоритма подразумевает построение структурной модели символа, составляющими которой являются вершины, изгибы, соединяющие и композитные рёбра. Каждая из структурных составляющих дополняется числовыми значениями для описания её относительного положения, размера и формы. Третья стадия предложенного алгоритма заключается в попарном сравнении структурной модели рассматриваемого символа со структурными моделями эталонных символов. В работе предложены три критерия схожести структурных моделей символов: на основе нахождения максимального паросочетания наименьшего веса, на основе метода пересечений и на основе вероятностного теста. Алгоритмы для вычисления всех предложенных критериев схожести структурных моделей символов имеют полиномиальную сложность, однако критерий на основе нахождения максимального паросочетания обладает лучшим быстродействием, обеспечивая более детальное сопоставление структурных моделей.

Кроме того, автором приводится алгоритм сегментации рукописного текста, идея которого заключается в построении предложенных структурных моделей не только для отдельных символов, но и для целых слов. В отличие от подавляющего большинства алгоритмов сегментации рукописных текстов, предложенный подход может разделять соседние символы с помощью невертикальных линий, не требуя обучения с применением лигатур. Алгоритм выбора линий для разделения соседних символов основан на идее динамического программирования. В работе приводится описание применяемой функции динамического программирования, а также оценка количества достижимых состояний динамического программирования.

В *третьей* главе приводится краткий обзор библиотек для реализации алгоритмов обработки изображений и методов компьютерной геометрии, после чего описывается программная реализация предложенных диссертантом алгоритмов в виде модульного приложения на основе выбранных библиотеки компьютерного зрения OpenCV и библиотеки геометрических вычислений Eigen. В приложении подробно описываются средства выбранного языка программирования и общеизвестных библиотек, использованные для написания программного приложения, а также приводятся диаграммы классов и пакетов, дополненные подробным описанием классов модуля распознавания символов и их методов.

В *четвёртой* главе приводятся результаты сравнения предложенных в диссертации алгоритмов с другими алгоритмами, способными функционировать в условиях малой обучающей выборки. Кроме того, экспериментально подтверждается необходимость применения алгоритма скелетизации Ву-Цая к результату алгоритма скелетизации Зонга-Суня и необходимость выполнения предварительной операции устранения плоских окончаний начертания перед процедурой скелетизации.

Далее в четвертой главе проводится исследование предложенного алгоритма сегментации рукописного текста и исследуется точность определения количества символов. В случаях, когда верно определено количество символов, оценивается точность классификации отдельных символов, дополнительно приводится распределение классифицируемых символов по порядковому номеру гипотезы, которая соответствует верному символу. Кроме того, приводится экспериментальное подтверждение оценки количества состояний динамического программирования при поиске оптимального множества линий, разделяющих отдельные символы сегментируемого слова.

В *заключении* приведены основные результаты работы.

В *приложении* приводятся акты внедрения и свидетельство о регистрации программы на ЭВМ.

Научная новизна работы

Основная научная новизна работы состоит в следующем:

1. Предложен алгоритм скелетизации бинарных изображений символов на основе комбинированного подхода с применением операции предварительного устранения плоских окончаний начертания символа и алгоритмов скелетизации Зонга-Суня и Ву-Цая, который обладает высоким быстродействием и позволяет получить скелетизированное представление начертания символа без удаления его важных элементов.

2. Предложена структурная модель символа, которая отличается от известных графовых моделей принципом разделения ключевых точек и изгибов, группировкой соединяющих рёбер в композитные, дополнительными метками точек и рёбер, и позволяет описать его топологию и геометрическую форму.

3. Предложен алгоритм построения структурной модели символа, позволяющий выделить структурные составляющие его начертания, который отличается от известных отсутствием необходимости применения дополнительных итераций алгоритма Ли для определения геометрических характеристик структурных составляющих.

4. Предложен критерий схожести структурных моделей символов, который отличается от аналогов идеей перехода от геометрического представления моделей к общеизвестной задаче нахождения максимального паросочетания наименьшего веса для существенного повышения точности распознавания символов в условиях малой обучающей выборки.

5. Предложен устойчивый по отношению к наклону символов алгоритм сегментации рукописного текста, который позволяет решать задачу сегментации текста в условиях малой обучающей выборки, не требуя для настройки использования изображений лигатур.

Степень обоснованности и достоверности результатов исследования

Все положения и выводы диссертации достоверны и научно обоснованы. Полученные результаты подтверждены проведёнными экспериментами на соответствующих наборах данных. Достоверность результатов работы подтверждается использованием стандартных способов оценки качества распознавания символов и согласованностью полученных результатов с результатами ранее известных алгоритмов.

Теоретическая значимость

Разработанные в диссертации алгоритмы, помимо рассмотренной в работе задачи распознавания рукописных символов, могут быть обобщены и адаптированы для применения при решении задач классификации отпечатков пальцев, идентификации почерка, проверки подписей на подлинность и большого класса задач анализа бинарных изображений.

Кроме того, описанная структурная модель символа может применяться в ранее предложенных алгоритмах распознавания символов, а

предложенный критерий схожести структурных моделей может быть использован для других типов моделей.

Практическая значимость

Описанные в работе алгоритмические и программные средства позволяют эффективно решать задачу распознавания рукописных символов в условиях малой обучающей выборки, когда применение универсальных подходов на основе признаков классификаторов существенно осложняется. Описанные условия возникают в процессе обработки изображений документов, требующем учёта индивидуальных особенностей почерка.

Рекомендации по использованию результатов работы

Результаты работы представляют существенный практический интерес. Они могут быть использованы в организациях, разрабатывающих системы распознавания рукописного текста, анализа особенностей почерка, идентификации подписи по единственному образцу и другие схожие системы. Рекомендуется использовать их в научной работе и при организации учебного процесса в образовательных учреждениях, занимающихся вопросами компьютерного зрения и компьютерной геометрии.

Замечания

1. В представленном аналитическом обзоре методов и алгоритмов распознавания символов не представлены алгоритмы полуавтоматического обучения, которые решают схожую задачу, хотя и используют элементы реальных тестовых данных для устранения недостатков обучения на малой обучающей выборке, чего не требуется предложенным в диссертации алгоритмам.

2. На стадии скелетизации бинарного изображения символа предлагается применять алгоритм Ву-Цая для устранения областей, для которых скелетизация выполнена не полностью. В разделе 4.5 приводится сравнение качества распознавания с применением алгоритма Ву-Цая с аналогичным показателем без его использования. Однако в работе нет подробного описания трудностей для алгоритма построения структурной модели символа, к которым приводят участки неполной скелетизации.

3. В работе исследуется применение критериев схожести, адаптированных специально для предложенной структурной модели символа, но не приводятся результаты для общеизвестных критериев, например, критерия на основе внутренних расстояний и критерия дистанционного преобразования.

4. Исследуемое количество состояний динамического программирования в процессе сегментации рукописного текста было бы

полезно сравнить с количеством состояний, которое бы было получено при полном переборе без применения кеширования, характерного для метода динамического программирования.

Заключение

Указанные замечания не влияют на общую положительную оценку работы. Диссертация выполнена на высоком уровне и обладает достаточной научной и практической значимостью. Результаты Хаустова П.А. в полной мере представлены в его публикациях и правильно отражены в автореферате. Автореферат полностью соответствует диссертационной работе.

Диссертационная работа «Алгоритмы распознавания рукописных символов в условиях малой обучающей выборки» удовлетворяет требованиям, предъявляемым к диссертациям на соискание учёной степени кандидата наук пунктом 9 действующего «Положения о присуждении учёных степеней», а её автор – Хаустов Павел Александрович – заслуживает присуждения ему ученой степени кандидата технических наук по специальности 05.13.11 – Математическое и программное обеспечение вычислительных машин, комплексов и компьютерных сетей.

Диссертация и отзыв на нее заслушаны, обсуждены и одобрены на заседании расширенного научно-технического семинара кафедры системного анализа и исследования операций института информатики и телекоммуникаций Сибирского государственного университета науки и технологий имени академика М.Ф. Решетнева, протокол № 4 от 29 ноября 2017 г.

Председатель семинара,
профессор кафедры САИО,
д-р техн. наук, профессор

Е.С. Семенкин

Семенкин Евгений Станиславович
Почтовый адрес: Россия, 660037, г. Красноярск,
пр. им. газ. «Красноярский рабочий», 31
Телефон: +7-391-291-91-41
E-mail: eugenesemenkin@yandex.ru
Сайт: <http://www.sibsau.ru>