

Министерство образования и науки Российской Федерации  
ФЕДЕРАЛЬНОЕ ГОСУДАРСТВЕННОЕ БЮДЖЕТНОЕ  
ОБРАЗОВАТЕЛЬНОЕ УЧРЕЖДЕНИЕ ВЫСШЕГО ОБРАЗОВАНИЯ  
«САРАТОВСКИЙ НАЦИОНАЛЬНЫЙ ИССЛЕДОВАТЕЛЬСКИЙ  
ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ ИМЕНИ  
Н.Г.ЧЕРНЫШЕВСКОГО»

Кафедра информатики и программирования

**ПРОГРАММИРОВАНИЕ КОМПЬЮТЕРНОГО ЗРЕНИЯ:  
РАСПОЗНАВАНИЕ ТЕКСТА НА ИЗОБРАЖЕНИЯХ СО СЛОЖНЫМ  
ФОНОМ**

АВТОРЕФЕРАТ МАГИСТЕРСКОЙ РАБОТЫ

студентки 2 курса 273 группы  
направления 01.04.02 Прикладная математика и информатика  
факультета компьютерных наук и информационных технологий  
Кустовой Евгении Олеговны

Научный руководитель:

к.ф.-м.н.

\_\_\_\_\_

М. В. Огнева

подпись, дата

Зав. кафедрой:

к.ф.-м.н.

\_\_\_\_\_

М. В. Огнева

подпись, дата

Саратов 2018

## ВВЕДЕНИЕ

**Актуальность темы.** Компьютерное зрение – это вид деятельности, в котором для извлечения данных применяются статистические методы и используются модели, построенные с помощью геометрии, физики и теории обучения. Компьютерное зрение основывается на четком представлении о камерах и физическом процессе формирования изображений (так как из изображений в дальнейшем извлекаются данные), получении простых выводов на основе изучения набора отдельных пикселей, умении суммировать информацию, полученную из множества изображений, упорядочении группы пикселей с целью их разделения или получения информации о форме, распознавании объектов с помощью геометрической информации или вероятностных методов. Компьютерное зрение применяется довольно широко как в относительно старых областях (например, управление мобильными роботами, средства наблюдения, военные приложения), так и в сравнительно новых (взаимодействие человек-компьютер, поиск изображений в цифровых библиотеках, анализ медицинских изображений и передача смоделированных сцен в компьютерной графике).

Отличительная черта компьютерного зрения – это извлечение данных из изображений или последовательности изображений. Это весьма полезная особенность. Процесс снятия изображения достаточно прост и на сегодняшний момент недорог. Данные, необходимые пользователям, могут в большой мере зависеть от области их применения. Например, такой аспект компьютерного зрения, как определение структуры по движению, позволяет из серии изображений получить представление о том, что изображено на рисунке и как движется камера. В индустрии развлечений подобные методы применяются для отсеивания движения и построения трехмерных компьютерных моделей с сохранением структуры. Эти модели применяются там, где нельзя использовать настоящие материалы (модели взрывают, поджигают и т.д.). С помощью небольшого числа фотографий можно получить хорошие, простые, удобные и точные модели. Есть еще целый ряд

важных областей применения компьютерного зрения. Это, например, работа с медицинскими изображениями: создание программных систем, которые могут улучшать набор изображений, выявлять на них важные моменты или события либо анализировать информацию, полученную из изображений. Другая важная область – различные технические проверки, когда по изображениям объектов определяется, соответствуют ли объекты спецификации. Третья сфера применения компьютерного зрения – интерпретация (раскрытие смысла) фотографий, сделанных со спутника, как в военных целях, так и в гражданских. На данный момент пользователь уже умеет находить нужную информацию в текстовых библиотеках, но еще не всегда знает, что делать с библиотеками статических и движущихся изображений.

Актуальность компьютерного зрения в современном мире и определила цель магистерской работы.

**Цель магистерской работы:** программирование компьютерного зрения: распознавание текста на изображениях со сложным фоном.

**Задачи магистерской работы:**

1. Определить понятие распознавания текста.
2. Изучить алгоритмы распознавания текста: оффлайн-распознавание и CSS алгоритм.
3. Сделать подбор изображений.
4. Предварительно подготовить изображения.
5. Реализовать алгоритм оффлайн-распознавания и CSS алгоритм.
6. Провести сравнительный анализ алгоритмов на наборе изображений.
7. Проанализировать результаты.

**Методологические основы** компьютерного зрения представлены в трудах Козлова [2], Балахонцевой [5], Грина [8], а также в книге Стокмана «Компьютерное зрение» [13].

**Научная значимость магистерской работы.** В ходе выполнения магистерской работы были выполнены все поставленные задачи, а именно

определено понятие распознавания текста, изучен алгоритм оффлайн-распознавания и CSS алгоритм, сделан подбор изображений для тестирования, реализована предварительная обработка изображений, алгоритм оффлайн-распознавания и CSS алгоритм, проведен сравнительный анализ алгоритмов на наборе изображений и проанализированы полученные результаты.

По итоговым результатам можно сказать, что CSS алгоритм и алгоритм оффлайн-распознавания хорошо подходят для изображений с монотонным фоном, для изображений с цветным фоном и черным текстом, где нет резких переходов, и менее подходят для полностью цветных изображений, где фон может сливаться с текстом и иметь резкие переходы. Стоит отметить, что CSS алгоритм можно использовать для словосочетаний и предложений.

**Структура и объем бакалаврской работы.** Работа состоит из введения, двух разделов, заключения, списка использованных источников информации и 7 приложений. Общий объем работы – 58 страниц, из них 54 страницы – основное содержание, включая 76 рисунков и 1 таблицу, список использованных источников информации – 28 наименований.

## КРАТКОЕ СОДЕРЖАНИЕ РАБОТЫ

**Первый раздел «Распознавание текста»** посвящен определению распознавания текста в целом и содержит несколько подразделов.

**Общая постановка задачи** включает в себя перечень условий и подготовительных этапов при распознавании текста.

**Основные подходы к распознаванию текста** – описывает три главных метода распознавания текста.

**Алгоритм оффлайн-распознавания** текста проводится по следующей схеме:

- 1) предобработка изображения, выделение области интереса;
- 2) сегментация и нормализация текста из области интереса;
- 3) распознавание сегментированного текста тем или иным методом.

**CSS алгоритм** – включает в себя основные этапы:

- 1) Получение контурного изображения детектором Кенни;
- 2) Морфологическое замыкание в случае, если буква имеет разрывы;
- 3) Выделение внешнего контура буквы;
- 4) Нормирование координат контура по длине дуги, получение плоской кривой;
- 5) Сглаживание кривой одномерной Гауссианой;
- 6) Расчет кривизны кривой вейвлетом Гаусса;
- 7) Получение CSS изображения.

**Детектор края Кенни** посвящен оператору Кенни, который выполняет контурный анализ на изображении, а также описывает шаги алгоритма.

**Вейвлеты Гаусса** описывают применение вейвлетов Гаусса для формулировки масштабно-пространственного представления кривизны контура замкнутой области.

**Аспекты практического применения** включает в себя построение модели масштабно-пространственного представления кривизны контура для букв латинского алфавита.

**Предварительная обработка изображений** производит улучшение качества изображения – приведение изображения к виду, наиболее подходящему для дальнейшей автоматической работы с ним. Улучшение качества изображения с целью распознавания текста включает в себя удаление дефектов изображения и отделение текста от фона.

**Сегментация изображений** описывает способы выделения букв на изображении, полученном на этапе предварительной обработки. Результатом этапа сегментации должна быть последовательность изображений отдельных букв, которые далее поступают на вход функции распознавания.

### **Второй раздел «Реализация алгоритмов и сравнительный анализ»**

Содержит реализации алгоритмов распознавания текста: алгоритма оффлайн-распознавания и CSS алгоритма, а также их сравнительный анализ на нескольких наборах изображений (простых, с цветным фоном и черным текстом и полностью цветных).

CSS алгоритм имеет меньшую погрешность при распознавании текста, чем алгоритм оффлайн-распознавания. Он очень хорошо подходит для простых изображений с монотонным фоном, а также может использоваться для изображений черным текстом и цветным фоном, где фон не имеет резких переходов. Кроме того, этот алгоритм способен распознавать несколько слов, в отличие от оффлайн-распознавания.

Алгоритм оффлайн-распознавания также может использоваться для простых изображений, только если они не содержат букв Ё, Ъ, Ы, и для изображений черным текстом и цветным фоном без резких переходов.

При этом оба алгоритма для полностью цветных изображений в случае, когда текст сильно сливается с фоном, дают неправильный результат.

## **ЗАКЛЮЧЕНИЕ**

В ходе выполнения магистерской работы были выполнены все поставленные задачи, а именно определено понятие распознавания текста, изучен алгоритм оффлайн-распознавания и CSS алгоритм, сделан подбор изображений для тестирования, реализована предварительная обработка изображений, алгоритм оффлайн-распознавания и CSS алгоритм, проведен сравнительный анализ алгоритмов на наборе изображений и проанализированы полученные результаты.

По итоговым результатам можно сказать, что CSS алгоритм и алгоритм оффлайн-распознавания хорошо подходят для изображений с монотонным фоном, для изображений с цветным фоном и черным текстом, где нет резких переходов, и менее подходят для полностью цветных изображений, где фон может сливаться с текстом и иметь резкие переходы. Стоит отметить, что CSS алгоритм можно использовать для словосочетаний и предложений.

## СПИСОК ИСПОЛЬЗОВАННЫХ ИСТОЧНИКОВ

1. Саратовский национальный исследовательский государственный университет имени Н.Г. Чернышевского [Электронный ресурс]. URL: <http://www.sgu.ru/> (дата обращения: 13.10.2017).
2. Козлов В.Д. Алгоритмы распознавания рукописных цифр [Электронный ресурс]. – Режим доступа: [http://www.machinelearning.ru/wiki/images/2/25/2015\\_417\\_KozlovVD.pdf](http://www.machinelearning.ru/wiki/images/2/25/2015_417_KozlovVD.pdf) . – Дата обращения: 4.12.2017
3. OCR System FineReader by ABBY [Электронный ресурс] – Режим доступа: <https://studfiles.net/preview/4444678/> . — Дата обращения: 5.12.2017
4. Сравнение систем оптического распознавания [Электронный ресурс] – Режим доступа: [http://masters.donntu.org/2012/fknt/gaydukov/library/article\\_3.htm](http://masters.donntu.org/2012/fknt/gaydukov/library/article_3.htm). — Дата обращения: 5.12.2017
5. Балахонцева А., Годоба А. //Система распознавания символов на изображениях со сложным фоном/ The 23rd International Conference on Computer Graphics and Vision. — 2013
6. Багрова И. А., Грицай А. А., Сорокин С. В., Пономарев С. А., Сытник Д. А. Выбор признаков для распознавания печатных кириллических символов // Вестник Тверского Государственного Университета 2010 г., 28, стр. 59-73
7. Kopf S., Haenselmann T., Effelsberg, W. Enhancing curvature scale space features for robust shape classification. Multimedia and Expo, 2005. ICME 2005. IEEE International Conference, 2005
8. Билл Грин. Алгоритм выделения контуров CANNY, Дрексельская лаборатория автоматизированных систем, 2002.
9. J. F. Canny, "Finding edges and lines in images," M.I.T. Artificial Intell. Lab., Cambridge, MA, Rep. AI-TR-720, 1983.
10. Оператор Собеля [Электронный ресурс] – Режим доступа: <https://www.intuit.ru/studies/courses/10621/1105/lecture/17989?page=6>. — Дата обращения: 31.03.2018

11. Хованова Н.А., Хованов И.А. Методы анализа временных рядов: Учеб. пособие. — Саратов: — Издательство ГосУНЦ Колледж. — 2001.
12. Mokhtarian F. and Mackworth A.K. A theory of multi-scale curvature based shape representation for planar curves // IEEE Transactions on Pattern Analysis and Machine Intelligence. — 1986. — Vol. 8, № 1. — P. 34-43.
13. Л Шапиро и Дж Стокман. «Компьютерное зрение». В: М.: Бином. Лаборатория знаний. — 752 (2006) . — с. 8.
14. Felipe C Ribas e A. Handwritten digit segmentation: a comparative study. A: International Journal on Document Analysis and Recognition (IJ DAR) 16.2 (2013), n. 127\_137.
15. Сакович И.О., Белов Ю.С. Обзор основных методов контурного анализа для выделения контуров движущихся объектов // Инженерный журнал: наука и инновации. — 2014. — вып. 12. URL: <http://engjournal.ru/catalog/it/hidden/1280.html>
16. Коробов Д. В., Патин М. В. Метод распознавания шрифта текста с изображения // Молодой ученый. — 2016. — №12. — С. 161-165. — URL <https://moluch.ru/archive/116/31528/> (дата обращения: 05.12.2017)
17. Mokhtarian F., Mackworth A. A Theory Multiscale, Curvature-Based Shape Representation for Planar Curves//IEEE Transactions on Pattern Analysis And Machine Intelligence, Vol. 14, No. 8, august 1992.
18. Морфологическая обработка изображений [Электронный ресурс]. — Режим доступа: <http://old.compsciclub.ru/node/1990>. — Дата обращения: 31.03.2018
19. Miscellaneous Image Transformations [Электронный ресурс]. — Режим доступа: [https://docs.opencv.org/2.4/modules/imgproc/doc/miscellaneous\\_transformations.html](https://docs.opencv.org/2.4/modules/imgproc/doc/miscellaneous_transformations.html) — Дата обращения: 28.05.2018
20. Козлов В.Д. Алгоритмы распознавания рукописных цифр [Электронный ресурс]. — Режим доступа:

Дата обращения: 28.05.2018

21. J. F. Canny, "Finding edges and lines in images," M.I.T. Artificial Intell. Lab., Cambridge, MA, Rep. AI-TR-720, 1983.

22. Morphological Transformations [Электронный ресурс]. – Режим доступа: [https://docs.opencv.org/trunk/d4/d76/tutorial\\_js\\_morphological\\_ops.html](https://docs.opencv.org/trunk/d4/d76/tutorial_js_morphological_ops.html). – Дата обращения: 06.04.2018

23. Гонсалес Р., Вудс Р. Цифровая обработка изображений. – Москва: Техносфера, 2005.

24. Скляренко Е.Г. Курс лекций по классической дифференциальной геометрии. – Москва, 2008.

25. Garcia G. B., Suarez O. D., Aranda J. L. E. Learning Image Processing with OpenCV. ISBN-13: 978–1-78328–765–9. Birmingham: Packt Publishing, 2015.

26. Люстерник Л.А., Лаврентьев М.А. Основы вариационного исчисления. Том 1. Часть 2. – ОНТИ, 2012.

27. Кустова Е. О. Визуализация задач аналитической геометрии с использованием системы Paraview // Информационные технологии в образовании: Материалы VIII Международ. научно-практ. конф. – Саратов: ООО «Издательский центр «Наука»», 2016. – 486 с.

28. Огнева М. В., Кустова Е. О. Технологии компьютерного зрения в научно-исследовательской работе студента: распознавание почерка // Информационные технологии в образовании: Материалы IX Международ. научно-практ. конф. – Саратов: ООО «Издательский центр «Наука»», 2017. – 486 с.