МИНИСТЕРСТВО НАУКИ И ВЫСШЕГО ОБРАЗОВАНИЯ РОССИЙСКОЙ ФЕДЕРАЦИИ

Федеральное государственное автономное образовательное учреждение высшего образования

**«КРЫМСКИЙ ФЕДЕРАЛЬНЫЙ УНИВЕРСИТЕТ имени В. И. ВЕРНАДСКОГО»**

(ФГАОУ ВО «КФУ им. В. И. Вернадского»)

**Таврическая академия (структурное подразделение)**

**Факультет математики и информатики**

**Кафедра информатики**

Иванча Николай Александрович

. **Разработка объектов дополненной реальности при посещении городов Крыма**

Выпускная квалификационная работа

Обучающегося 4 курса

Направления подготовки 01.03.02. Прикладная математика и информатика

Форма обучения очная

Научный руководитель

Доцент кафедры информатики,

Кандидат физико-математических наук М. Г. Козлова

К ЗАЩИТЕ ДОПУСКАЮ:

Заведующий кафедрой,

доктор физико-математических наук,

профессор В. И. Донской

Симферополь, 2019

**Аннотация**

**Иванча Н. А. Разработка объектов дополненной реальности при посещении городов Крыма.** Выпускная квалификационная работа бакалавра по направлению 01.03.02. Прикладная математика и информатика. Таврическая академия Крымского федерального университета имени В. И. Вернадского.

Целью работы является создание приложения дополненной реальности, использую оптимальное программное обеспечение. Содержание работы состоит в исследовании, тестировании и разборе современных технологий, позволяющих работать с дополненной реальностью. Результатом работы является создание мобильного приложения.

Ключевые слова: дополненная реальность, AR, распознавание, текст, оптические системы распознавания, OCR, бинаризация, макет, маркер, target, объект, модель, моделирование, редактор, программное обеспечение, среда разработки, предобработка, шум, искажение, изображение, методы, классификация, шаг, сегментация, обучение, анализ, приложение, визуализация, город, Крым.

Страниц – 55, таблиц – 0, иллюстраций – 24, приложений – 1, библиографических источников – 37.

Содержание

[Введение 4](#_Toc10713846)

[1. Подходы к распознаванию текста 7](#_Toc10713847)

[1.1. Оптическое распознавание символов 7](#_Toc10713848)

[1.2. Метод распознавания текста №1 11](#_Toc10713849)

[1.3. Метод распознавания текста №2 13](#_Toc10713850)

[1.4. Метод распознавания текста №3 14](#_Toc10713851)

[1.5. Распознавание текста по структуре скелета букв 16](#_Toc10713852)

[1.6. Методы бинаризации изображений 19](#_Toc10713853)

[1.7. Tesseract OCR 21](#_Toc10713854)

[2. Подходы к работе с объектами дополненной реальности 29](#_Toc10713855)

[2.1. Vuforia 32](#_Toc10713856)

[2.2. ARToolKit 34](#_Toc10713857)

[2.3. Подходы к разработке объектов дополненной реальности 35](#_Toc10713858)

[2.3.1. Blender 35](#_Toc10713859)

[2.3.2. 3ds Max 36](#_Toc10713860)

[2.3.3. ZBrush 37](#_Toc10713861)

[3. Разработка комплекса визуализации объектов дополненной реальности при посещении городов Крыма 38](#_Toc10713862)

[3.1. Распознавание текста 38](#_Toc10713863)

[3.2. Создание объектов дополненной реальности 38](#_Toc10713864)

[3.3. Визуализация элементов дополненной реальности 38](#_Toc10713865)

[Приложение 48](#_Toc10713866)

[Список источников 53](#_Toc10713867)

# Введение

В последние годы несложно заметить стремительное развитие технологий, цель которых облегчить деятельность человека. Одной из таких технологий является augmented reality (AR) или дополненная реальность. Дополненная реальность – это технология, которая в реальном времени дополняет привычный для человека физический мир, элементами компьютерной графики (картинки, текст и др.), используя при этом специальное программное обеспечение и доступные каждому – смартфоны, планшеты или другие подобные устройства.

Очень важно отличать дополненную реальность от виртуальной (virtual reality, VR) и смешанной (mixed reality, MR). В отличие от дополненной реальности, где виртуальные объекты проецируются на реальное окружение, виртуальная представляет собой новый, созданный техническими средствами мир, который передаётся человеку через органы зрения, слуха и осязания. Смешанная реальность является наиболее сложной, так как объединяет оба подхода.

Дополненная реальность уверенно занимает лидирующие позиции на рынке технологий. Это подтверждает столь частое упоминание AR в презентациях тех или иных IT-гигантов. Такой интерес к дополненной реальности вызван тем, что она предоставляет возможность взаимодействовать с окружением на принципиально новом уровне, а также имеет огромный спектр областей применения.

Основным двигателем AR технологий на данный момент является индустрия развлечений, которая видит в дополненной реальности возможность установления более тесных связей между аудиторией и своими персонажами. Так, например, в конце 2016 года компанией Niantic была выпущена мобильная игра про «карманных монстров» Pokemon GO, которая всего за несколько месяцев побила рекорды популярности среди мобильных приложений и положила начало кардинальным изменениям в индустрии развлечений. [1]

Не обошли стороной новые технологии и рекламную индустрию. С появлением дополненной реальности различные компании открыли для себя новые способы донесения привлекательности своего продукта до своих потенциальных потребителей. Дополненную реальность используют для наполнения плоских поверхностей листовок и билбордов интерактивным контентом, который имеет огромное преимущество перед статичной рекламой. Например, известный журнал об автомобилях Top Gear предоставил своим читателям возможность просматривать видеоролики об автомобилях прям на страницах издания, используя для этого только лишь смартфон со специальным приложением. Такой же рекламный ход совершила и компания Faberlic, предоставив своим клиентам возможность посмотреть на косметические товары из своего печатного каталога в натуральную величину, прочитать состав на этикетке и сделать заказ. [2]

Несмотря на огромную популярность дополненной реальности в сферах рекламы и развлечения, потенциал использования AR намного шире, что наглядно подтверждают современные примеры применения технологии в медицине. Первый самый очевидный пример использования AR в медицине – это организация максимально наглядного обучения студентов медицинских вузов. Наглядную анатомию в дополненной реальности способно продемонстрировать приложение HoloAnatomy, разработанное компанией Microsoft для очков смешанной реальности HoloLens. [3]

Более серьёзными с точки зрения медицины являются приложения: MITK pille, которое позволяет, используя планшет, прямо во время операции «заглянуть» в тело пациента, и приложение для людей с ограниченными возможностями зрения и слуха Aira, в котором реализован нейросетевой помощник, распознающий и проговаривающий всё, что видит камера очков. [4]

Своё применение дополненная реальность нашла и в сфере дизайна и проектирования. Так, например, компания Ikea интегрировала AR в своё приложение Ikea Place, с помощью которого позволяет пользователям проверить, как мебель может выглядеть в их домах.

Также дополненная реальность оказалась востребована в сфере туризма. Уже существуют приложения, помогающие людям, оказавшимся в незнакомом городе, успешно ориентироваться на местности и прокладывать маршруты к необходимым локациям. Примером такого приложения является разработка компании MapWay – приложение Bus Times London, которое, используя камеру смартфона, помогает туристам найти свою автобусную остановку в пределах города Лондон, рисуя средствами дополненной реальности путь до неё.

При посещении музеев, памятников архитектуры или других достопримечательностей средствами дополненной реальности могут быть реализованы такие опции как виртуальный экскурсовод или визуализация каких-либо событий, связанных с данной достопримечательностью.

Технология дополненной реальности востребована в столь многих сферах деятельности человека по следующим причинам:

1. Возможность получать информацию в реальном времени.
2. Необычный способ представления информации, благодаря которому возможно привлечь больше внимания окружающих и более точно донести до них необходимую информацию.
3. Усиливает запоминание деталей и понимание процесса за счёт наглядности. Актуально для сферы образования.
4. Инновационность. Несмотря на уверенное положение на рынке технологий, воспринимается как нечто новое и современное.
5. Большой, вечно расширяющийся, спектр сфер использования.

Учитывая высокую популярность AR технологий, отличным решением будет внедрение элементов дополненной реальности в самую востребованную отрасль крымского полуострова – туризм. С помощью дополненной реальности можно визуализировать достопримечательности Крыма на физической карте, с целью привлечения большего числа туристов с разных уголков земного шара.

Целью работы является разработка программного обеспечения

1. **Подходы к распознаванию текста**

Следует отметить, что задача распознавания текста важна до сих пор даже при большом количестве уже готовых программ. Большое количество задач требует компактных алгоритмов, заточенных под решение конкретно требуемых проблем.

Распознавание текста является оной из важнейших подразделов распознавания образов. Существует большое количество литературы с огромным количеством информации, которая внесла неизмеримый вклад в развитие человечества, которую уже попросту нельзя взять в руки, не повредив книгу. Именно поэтому оцифровка такой литературы столь важна.

Следует отметить, что под распознаванием текста понимается три основных метода:

1. сравнение с подготовленными шаблонами;
2. распознавание, которое использует критерии распознаваемого объекта;
3. распознавание при помощи нейронных сетей, самообучающихся алгоритмов. [4]
   1. **Оптическое распознавание символов**

Оптическое распознавание символов – это механическое или электронное конвертирование текста на изображении в текст машинного кодирования.

Принцип работы оптического распознавания символов заключается в следующих этапах.

1. Процесс предобработки.

Оптические системы распознавания часто прибегают к предобработке изображения с целью улучшения результатов. Существуют такие виды предобработки: выравнивание (de-skew) – это ситуация, когда сфотографированный или отсканированный текст располагается под наклоном, когда нам необходим строго горизонтальный или вертикальный текст; удаление шумов (despeckle) – удаление пятен, сглаживание углов; бинаризация (binarisation) – перевод изображения из цветного в черно-белое; удаление линий (line removal) – удаление нерельефных полей и строк; анализ макета или «зонирование» (Layout analysis or «zoning») – определение столбцов, абзацев, подписей и т. д. как отдельные блоки, это особенно важно в многостолбцовых раскладках и таблицах; обнаружение строк и слов (line and word detection) – установка основной линии для слов и форм символа, разделение слова если необходимо; выделение символов или сегментация (character isolation or «segmentation») – текст изображения разделяется на символы; нормализация пропорций и масштаба изображений (normalize aspect ratio and scale).

2. Распознавание символов.

Здесь присутствуют два основных типа оптического распознавания.

Матричное сопоставление включает в себя сравнение изображения с сохраненным образом на попиксельной основе; это так же известно как «сопоставление с образом», «распознавание образов» или «корелляция изображения». Это основано на том, что входной знак правильно изолирован от остальной части изображения, а сохраненный рельеф имеет аналогичный шрифт того же масштаба. Этот метод работает лучше всего с машинописным текстом и не работает, когда встречаются новые шрифты. Это метод, который применялся в раннем физическом оптическом распознавании на основе фотоэлементов.

Извлечение объектов разбивает знаки на «объекты», такие как линии, замкнутые циклы, направление линий и пересечения линий. Особенности извлечения уменьшают размерность представления и делают процесс распознавания вычислительно эффективным. Эти функции сравниваются с абстрактным векторным представлением символа, которое может привести к одному или нескольким прототипам знаков. Общие методы обнаружения признаков в компьютерном зрении применимы к этому типу OCR, что обычно наблюдается в «интеллектуальном» распознавании рукописного ввода и, действительно, в большинстве современных программ OCR. Классификаторы ближайших соседей, такие как алгоритм k-ближайших соседей, используются для сравнения объектов изображения с сохраненными объектами знака и выбора ближайшего совпадения.

Программное обеспечение, такое как Cuneiform и Tesseract, использует двухфазный подход к распознаванию символов. Второй проход известен как «адаптивное распознавание» и использует формы букв, распознаваемые с высокой достоверностью на первом проходе, для лучшего распознавания оставшихся букв на втором проходе. Это выгодно для необычных шрифтов или сканов низкого качества, где шрифт искажен (например, размыт или выцветает).

3. Постобработка.

Точность распознавания можно повысить, если вывод ограничен лексиконом – списком слов, которые могут встречаться в документе. Это могут быть, например, все слова на английском языке или более технический словарь для конкретной области. Этот метод может быть проблематичным, если документ содержит слова, не входящие в лексикон, например, имена собственные. Tesseract использует свой словарь, чтобы повлиять на шаг сегментации символов, для повышения точности. Выходной поток может быть потоком простого текста или файлом символов, но более сложные системы распознавания могут сохранить исходный макет страницы и создать, например, аннотированный PDF-файл, который включает как исходное изображение страницы, так и текстовое представление с возможностью поиска. «Анализ ближнего соседа» может использовать частоты совпадений для исправления ошибок, отмечая, что некоторые слова часто встречаются вместе. Например, «Вашингтон, округ Колумбия» как правило, гораздо чаще встречается в английском, чем «Вашингтон DOC». Знание грамматики сканируемого языка может также помочь определить, может ли слово быть, например, глаголом или существительным, что позволяет повысить точность.

4. Оптимизация под конкретные приложения.

Основные поставщики технологий OCR начали настраивать системы OCR, чтобы лучше справляться с определенными типами ввода. Помимо лексики для конкретного приложения, лучшую производительность можно достичь, принимая во внимание правила, стандартные выражения, или богатую информацию, содержащуюся в цветных изображениях. Эта стратегия называется «Ориентированное на приложение OCR» или «Настраиваемое OCR» и применяется к распознаванию номеров, счетов, снимков экрана, удостоверений личности, водительских прав и производства автомобилей. То есть для улучшения распознавания определяются параметры, которые делают задачу более узкоспециализированной. [5]

Рассмотрим основные трудности при распознавании символов:

* разнообразие языков, символов, и способов изображать их. В некоторых языках присутствуют схожие символы: «N» и «И», «R» и «Я» и др. В некоторых языках присутствуют символы, которые схожи между собой или похожи на цифры. Например, в русском языке: «М», «Н», «И» и «О» и «0», и т.д.;
* искажение изображений символов. Причинами могут представляться огромное количество факторов. От незначительных проблем связанных изображением исходного формата, заканчивая дефектом самого источника;
* при обработке персональных данных, трудности могут возникать из-за методов защиты информации. Например, на паспортах присутствуют невидимые символы;
* так же усложняет задачу большое разнообразие различных шрифтов разных вариаций, размеров и масштабов.

Искажения цифровых изображений текстовых символов могут быть вызваны следующими причинами:

* шумами печати, в частности, не достаточно хорошей печатью (разрывами слитных черт символов), «слипанием» соседних символов, пятнами и ложными точками на фоне вблизи символов и т. п.;
* смещением символов или частей символов относительно их ожидаемого положения в строке;
* изменением наклона символов;
* искажением формы символа за счет оцифровки изображения с низкой степенью детализации;
* эффектами освещения (тени, блики и т. п.) при съемке камерой.

Большинство OCR-систем показывают высокие результаты в случае обработки черного текста на белом фоне. А при анализе не черно-белых изображений полученные данные могут быть некорректны.

Многие проблемы распознавания текста на цветных изображениях можно устранить на этапе предобработки. [6]

OCR-системы должны уметь обозначить на цифровом изображении области, в которых содержится текст, определять в них отдельные строки, выделять символы, распознавать их и быть устойчивыми к различным особенностям печати, расстоянию между символами, строками и т. д.

После изучения трех наиболее популярных приложений по распознаванию символов: ABBYY FineReader, CuneiForm, OCRopus, были сделаны следующие выводы.

1. Большинство из приложений обладают достаточно сложным интерфейсом и большим числом функций, которые рядовому пользователю могут быть попросту не нужны.
2. Все приложения (кроме FineReader) не обладают способностью работы в фоновом режиме.
3. Не все обладают способностью предобработки изображения.
4. Отсутствие графического интерфейса некоторых приложений.
5. Среди них нет программ, которые показывают идеальное распознавание текста. [7]

**Заключение**

В результате работы было разработано приложение, которое хорошо работает на черно-белых картах. Для усовершенствования предстоит включить в модуль, который будет заниматься предобработкой изображения, что значительно повысит точность распознавания названий городов и на цветных картах. Так же предстоит расширить спектр городов, достопримечательности которых так же будут находиться на карте. Некоторые города имеют несколько характерных достопримечательностей, из этого следует, что предстоит разработка нескольких 3D-моделей города и будущий интерфейс пользователя может получать информацию о типах достопримечательностей, которые хочет посетить пользователь и предоставлять ему только те, которые соответствуют данной категории, а если в городе несколько достопримечательностей одной категории, то будет возможен их поочередный просмотр.

Технология дополненной реальности в применении к туристической отрасли имеет огромный спектр возможностей. Окончательным вариантом приложения будет программа, которая, кроме визуализации достопримечательностей, будет способна проложить путь к необходимому месту с возможностью выбора способа передвижения пользователя и провести на нем экскурсию.

У модуля оптического распознавания Tesseract есть недостаток – это относится к поддержке русского языка, она есть, но данные обучены недостаточно в сравнении с интернациональными языками. Но при необходимости данные обучения можно создать и самостоятельно. Или через некоторое время разработчики могут их усовершенствовать.

# Приложение

**Список источников**

* + - 1. Мир после Pokemon Go: как технология дополненной реальности изменит нашу жизнь [Электронный ресурс] – Режим доступа: <https://www.bbc.com/russian/features-36898771/>.
      2. Дополненная реальность в маркетинге: 15+ примеров использования [Электронный ресурс] – Режим доступа: <https://texterra.ru/blog/dopolnennaya-realnost-v-marketinge-primery-ispolzovaniya.html>.
      3. HoloAnatomy [Электронный ресурс] – Режим доступа: <https://www.microsoft.com/ru-ru/p/holoanatomy/9nblggh4ntd3/>.
      4. VR и AR в медицине [Электронный ресурс] – Режим доступа: <https://zen.yandex.ru/media/id/5ad5fbd76104934709d6f592/vr-i-ar-v-medicine-5b17aa5d83090530bc94c7c4/>.
      5. Say hej to IKEA Place [Электронный ресурс] – Режим доступа: https://www.ikea.com/au/en/customer-service/mobile-apps/say-hej-to-ikea-place-pub1f8af050
      6. Виртуальная и дополненная реальность в медицине [Электронный ресурс] – Режим доступа: URL: https://make-3d.ru/articles/virtualnaya-i-dopolnennaya-realnost-v-medicine/.
      7. Мобильный гид с дополненной реальностью [Электронный ресурс] – Режим доступа: URL: https://4pda.ru/2018/01/15/349175/.
      8. Optical character recognition [Электронный ресурс] – Режим доступа: URL: [https://en.wikipedia.org/wiki/Optical\_character\_recognition#Techniques](https://en.wikipedia.org/wiki/Optical_character_recognition" \l "Techniques).
      9. NicomSoft [Электронный ресурс] – Режим доступа: URL: <https://www.nicomsoft.com/optical-character-recognition-ocr-how-it-works/>.
      10. Методы распознавания текста [Электронный ресурс] – Режим доступа: URL: <https://habr.com/ru/post/220077/>.
      11. Распознаем текст, используя расстояние Хэмминга [Электронный ресурс] – Режим доступа: URL: <https://habr.com/ru/post/90867/> .
      12. Декодирование капчи на Python [Электронный ресурс] – Режим доступа: URL: <https://habr.com/ru/post/149091/> .
      13. Разве Tesseract распознаёт медленно? [Электронный ресурс] – Режим доступа: URL: <https://habr.com/ru/company/smartengines/blog/300990/>.
      14. Лапкина, А. Л. Распознавание текста по структуре скелета букв/ А. Л. Лапкина. – М. 2018. – С. 3 – 19.
      15. Текстовые капчи легко распознаются нейронными сетями глубокого обучения [Электронный ресурс] – Режим доступа: URL: <https://habr.com/ru/post/343222/>.
      16. Пешков, А. О. Анализ методов бинаризации изображений / А. О. Пешков, В. К. Крят. – Ростов-на-Дону. 2015. – С. 2 – 7.
      17. Smith, R. An Overview of the Tesseract OCR Engine / – С. 1 – 5.
      18. Quora [Электронный ресурс] – Режим доступа: URL: <https://www.quora.com/How-does-the-Tesseract-API-for-OCR-work>.
      19. HackerNoon [Электронный ресурс] – Режим доступа: URL: <https://hackernoon.com/predictions-for-the-future-of-augmented-reality-63c7b8c9d794>.
      20. Смолин, А. А. Системы виртуальной, дополненной и смешанной реальности / А. А. Смолин, Д. Д. Жданов, И. С. Потемин, А. В. Меженин, В. А. Богатырев. – СПб. 2018. – С. 45 – 54.
      21. Funreality [Электронный ресурс] – Режим доступа: URL: <https://funreality.ru/technology/augmented_reality/>.
      22. Шакиров, И. Ш. «Дополненная реальность»: инновационная технология организации образовательного процесса по информатике / И. Ш. Шакиров. – Лесосибирск. 2016. – С. 10 – 36.
      23. Константинова, К. Э. Исследование возможностей и опыта использования технологий дополненной реальности в экспозиционно-выставочной деятельности музеев / К. Э. Константинова. – М. 2018. – С. 6 – 13.
      24. Vuforia: немного магии в нашей реальности [Электронный ресурс] – Режим доступа: URL: <https://habr.com/ru/post/198862/>.
      25. Dgng [Электронный ресурс] – Режим доступа: URL: <http://dgng.pstu.ru/conf2017/papers/125/>.
      26. Методы распознавания текстов [Электронный ресурс] – Режим доступа: URL: <https://habr.com/ru/post/112442/>.
      27. 3D редакторы, плюсы и минусы [Электронный ресурс] – Режим доступа: URL: <https://habr.com/ru/post/136350/>.
      28. Оптическое распознавание символов в Linux [Электронный ресурс] – Режим доступа: URL: <https://habr.com/ru/post/153617/>.
      29. Бинаризация изображений: алгоритм Брэдли [Электронный ресурс] – Режим доступа: URL: <https://habr.com/ru/post/278435/>.
      30. Сегментация изображения [Электронный ресурс] – Режим доступа: URL: <https://habr.com/ru/post/128768/>.
      31. **What's OCR?** [Электронный ресурс] – Режим доступа: URL: <http://www.dataid.com/aboutocr.htm>.
      32. Использование Open Source OCR библиотеки Tesseract в Android на примере простого приложения [Электронный ресурс] – Режим доступа: URL: <https://habr.com/ru/post/282582/>.
      33. Как создать приложение дополненной реальности с помощью ARCore [Электронный ресурс] – Режим доступа: URL: <https://habr.com/ru/post/439190/>.
      34. Что такое ARCore? Всё, что вам нужно знать [Электронный ресурс] – Режим доступа: URL: <https://habr.com/ru/post/437378/>.
      35. Создание вашего первого ARCore-приложения [Электронный ресурс] – Режим доступа: URL: <https://habr.com/ru/post/438178/>.
      36. Создание AR-игры с помощью Vuforia [Электронный ресурс] – Режим доступа: URL: <https://habr.com/ru/post/440592/>.
      37. Azoft [Электронный ресурс] – Режим доступа: URL: http://www.azoft.ru/ blog/biblioteki-dopolnennoj-realnosti-dlya-ar-prilozhenij/.
      38. Афонасенко, А. В. Обзор методов распознавания структурированных символов / А. И. Елизаров. – М.: Технические науки, 2008. – С. 83 – 87.
      39. Вахрушева, Т. С. Применение технологий дополненной реальности в образовании / Т. С. Вахрушева. – СПб. 2017. – С. 33 – 40.
      40. Меженин, В. А. Богатырев. – СПб. 2018. – С. 45 – 54.
      41. GOLOS [Электронный ресурс] – Режим доступа: URL: <https://golos.io/vox-populi/@vp-webdev/raspoznavanie-teksta-na-izobrazhenii>.