**Доклад**

Слайд 2

Мною была подготовлена выпускная квалификационная работа на тему **Разработка объектов дополненной реальности для визуализации туристических объектов Крыма.**

**Целью работы** является разработка программного обеспечения реализующего визуализацию туристических объектов Крыма в дополненной реальности.

**Задачи**, решаемые в работе:

1. обзор научных источников по данному вопросу;
2. изучение необходимых программных средств и библиотек;
3. создание 3D-моделей достопримечательностей;
4. реализация программного продукта для распознавания названий городов на изображении (карте);
5. реализация программного продукта для визуализации 3D-моделей достопримечательностей.

Слайд 3

В ходе подготовки выпускной работы были рассмотрены этапы создания мобильного приложения дополненной реальности с распознаванием текста. Вся работа была разделена на две центральные проблемы: **Распознавание текста** и **Дополненная реальность**.

Работа над распознаванием текста предполагала решение следующих задач:

* захват кадра (изображения)
* предобработка изображения
* распознавание символов
* нечёткое сравнение с шаблоном
* постобработка текста

При работе над дополненной реальностью были решены следующие задачи:

* создание 3D моделей достопримечательностей
* предобработка изображения
* распознавание целевой поверхности
* оценка гомографии
* проекция 3D модели на целевую поверхность

Слайд 4

Процесс распознавания текста может быть реализован при помощи четырёх основных методов:

1. **Шаблонные методы** – сравнение распознанного объекта с подготовленными шаблонами.

В таких методах необходимо использовать метрику – некое условное значение функции, определяющее положение объекта в пространстве. Если два объекта близко друг к другу, то метрики для таких объектов будут совпадать или быть предельно похожими.

В качестве метрики можно выбрать расстояние Хэмминга, его часто используют при кодировании информации и передачи данных, оно показывает, как сильно объекты не похожи между собой. Таким образом, чтобы определить какая буква изображена на кадре, нужно найти её метрику со всеми готовыми шаблонами. И тот шаблон, расстояние Хэмминга до которого окажется наименьшим, и будет ответом.

1. **Признаковые методы** – распознавание, которое использует критерии распознаваемого объекта.

Эти методы являются наиболее распространёнными, признак в них определяется как функция от значений, содержащихся в одном или более пикселях, и вычисляется так, что численно выражает некоторую значимую характеристику объекта.

1. **Структурные методы** – распознавание, основанное на структуре объекта.

Такие методы переводят символ в его топологическое представление, отражающее информацию о взаимном расположении структурных элементов. В качестве структурных единиц могут выступать: линии, пятна, дуги окружностей, углы и т. д.

1. **Нейросетевые методы** – распознавание при помощи нейронных сетей, самообучающихся алгоритмов.

Нейрон, как структурная единица нейронной сети, получает множество входных сигналов. В нашем случае входные сигналы описывают значение пикселя изображения, то есть, если имеется изображение 16х16, входных сигналов у нейрона должно быть 256.

Количество выходных сигналов будет ограничено алфавитом, в нашем случае их 33, без учёта цифр.

Слайд 5

При подготовке к реализации программного продукта, было проведено исследование, тестирование и разбор современных технологий, позволяющих работать с распознаванием текста.

Для решения этой задачи, в итоге, был выбран подход с использованием библиотеки Tesseract. Tesseract обладает множеством преимуществ, опираясь на которые эта оптическая система и была выбрана:

1. обладает открытым исходным кодом;
2. показывает высокие результаты на монохромных изображениях;
3. высокая скорость и простота реализации задачи распознавания текста;
4. имеет возможность работать с библиотеками компьютерного зрения (OpenCV);
5. поддерживает интеграцию в большое количество языков программирования;
6. не коммерческий продукт.

Слайд 6

Алгоритм работы Tesseract довольно прост, его можно представить следующей схемой.

Верхняя строка – это встроенная в Tesseract предобработка изображения, которая предусматривает бинаризацию изображения и поиск контуров, для отсечения фона. Сам же процесс распознавания можно разделить на следующие этапы:

1. **Поиск строк.** Ключевыми частями этого этапа являются фильтрация объектов изображения и построение линий. Простой фильтр высоты удаляет выпадающие элементы и вертикально соприкасающиеся символы. Средняя высота приближается к размеру текста во всём слове, поэтому можно безопасно отсекать все те объекты, которые меньше средней высоты (знаки препинания, диакритические знаки и шум).
2. **Выбор базовой линии.** После того как строки текста были найдены, устанавливаются ограничительные линии.
3. **Поиск слов.** Данный этап нам не понадобится, так как текст, который мы хотим распознать – это одно слово – название города. С помощью определённых параметров для Tesseract`а можно пропустить этот этап.
4. **Определение фиксированного шага и разделение слова на буквы.** Tesseract проверяет текстовые строки, чтобы определить, как в них распределены буквы. Если буквы расположены с фиксированным шагом, то Tesseract разбивает это слово на символы.
5. **Распознавание слов.** В первую очередь классифицируется исходная сегментация, полученная в результате линейного поиска. Остальная часть распознавания на этом этапе применяется только к тексту без фиксированного шага. То есть, сначала, буквы «С», «и», «л», «ь» будут успешно распознаны, а после определится слово с новым разбиением.
6. **Разделение «слипшихся» символов и распознавание дефектных символов.**
7. **Классификация.** Протекает как двухэтапный процесс. На первом этапе классификатор создает короткий список классов, к которым может принадлежать неизвестный символ. На втором этапе для каждого признака неизвестного символа ищется битовый вектор прототипов данного класса, которому он может соответствовать, а затем вычисляется фактическое сходство между ними.

Слайд 7

Для моделирования 3D моделей достопримечательностей Крыма был выбран 3D-редактор 3ds Max, так как он имеет большие преимущества перед другими инструментами по части создания моделей архитектурных сооружений.

Распознавание AR маркеров и выведение 3D модели на опорную поверхность проводилось без использования готовых инструментов по следующим причинам:

1. отсутствие у представленных решений полностью бесплатных пакетов;
2. личная заинтересованность в самостоятельной реализации всех этапов визуализации 3D моделей в дополненной реальности;
3. отсутствие материально-технической базы для использования профессиональных продуктов для создания AR приложений.

Слайд 8