

Отчет о лабораторной работе № 3 «Обработка и распознавание изображений».

317 группа, ММП ВМК МГУ.

Булкин Антон Павлович.

26 мая 2025 г.

Содержание

1	Постановка задачи	2
2	Описание данных	2
3	Описание используемых методов	3
3.1	Кластеризация методом k-средних	3
3.2	Морфологическая обработка	3
3.3	Скелетизация и построение графа	3
3.4	Определение кончиков пальцев	4
3.5	Сортировка и определение состояния пальцев	4
3.6	Пороговая настройка параметров	4
3.7	Графический интерфейс	4
4	Описание программной реализации	4
4.1	Общая структура реализации	5
4.2	Чтение и предобработка изображений	5
4.3	Сегментация ладони	5
4.4	Скелетизация и построение графа	5
4.5	Определение кончиков пальцев	5
4.6	Распознавание позы руки	6
4.7	Графический интерфейс	6
4.8	Формат сохранения результатов	6
5	Демонстрация работы программы	7

1 Постановка задачи

В рамках лабораторной работы №3 требуется разработать программное решение, позволяющее классифицировать изображения ладоней с целью определения позы и линии пальцев. Задача предусматривает следующие этапы реализации:

- Ввод и отображение изображений ладоней, предоставленных в формате TIF.
- Сегментация изображений на основе точечных и пространственных преобразований.
- Определение позы ладони по расположению сомкнутых пальцев.
- Определение линии пальцев — ломаной линии, соединяющей кончики пальцев с точками в их основаниях.

Результаты работы программы должны быть представлены в виде текстового файла Results.txt, в котором указывается код позы ладони и координаты вершин линии пальцев.

Задача разбита на два уровня сложности:

- Intermediate: определение позы ладони и визуализация результата «изображение + код позы».
- Expert: дополнительно к уровню Intermediate — определение и визуализация линии пальцев.

В данной работе реализован уровень Intermediate.

2 Описание данных

Для выполнения работы предоставлен набор данных из 67 цветных изображений ладоней разных людей, полученных с помощью сканера. Изображения имеют размер 489×684 пикселей и разрешение 72 dpi. На изображениях представлены сканы левых рук.

Поза ладони определяется по расположению пальцев, которые нумеруются от 1 (большой палец) до 5 против часовой стрелки. Поза описывается кодом вида:

$$1 * 2 * 3 * 4 * 5$$

где «*» принимает значение «+», если пальцы прижаты друг к другу, и «», если пальцы разомкнуты.

В работе используются исходные изображения формата TIF и генерируется файл с результатами анализа, соответствующий формату, требуемому заданием.

3 Описание используемых методов

В данной работе для сегментации, анализа формы и определения состояния ладони были использованы методы компьютерного зрения, обработки изображений и анализа графов.

3.1 Кластеризация методом k-средних

Для начальной сегментации изображения ладони применялся метод кластеризации k-means. Метод позволяет разделить пиксели изображения на кластеры с похожими цветовыми характеристиками. Для изображения в цветовом пространстве RGB формулируется задача оптимизации:

$$\min_C \sum_{i=1}^k \sum_{x \in C_i} \|x - \mu_i\|^2, \quad (1)$$

где C_i — множества кластеров, μ_i — центры кластеров.

После выполнения кластеризации выбирается наиболее яркий кластер, соответствующий ладони.

3.2 Морфологическая обработка

Для очистки полученной бинарной маски ладони от шумов и артефактов применялись морфологические операции:

- **Эрозия** — уменьшает области и удаляет мелкие детали.
- **Дилатация** — увеличивает области и сглаживает контуры.

Последовательное применение этих операций улучшает качество сегментации, выделяя основную компоненту, соответствующую ладони.

3.3 Скелетизация и построение графа

Для анализа структуры ладони используется процедура скелетизации — преобразования бинарной маски в тонкий скелет, сохраняющий её топологические

свойства. Полученный скелет представляется в виде графа, где узлами являются пиксели скелета, а ребрами — соседство пикселей.

3.4 Определение кончиков пальцев

Для поиска кончиков пальцев в графе скелета выделяются узлы с одной степенью связности. Затем, используя геометрический анализ (кластеризация по углу и удаленности от центра ладони), выделяются 5 основных кончиков пальцев.

3.5 Сортировка и определение состояния пальцев

Пальцы сортируются с использованием алгоритма выпуклой оболочки (Convex Hull), начиная с большого пальца и далее в порядке указательный, средний, безымянный и мизинец. Для определения положения пальцев (сомкнуты или разомкнуты) используются расстояния между соседними кончиками пальцев. Пороговые значения для принятия решений были подобраны автоматически на основе данных.

3.6 Пороговая настройка параметров

Пороговые значения, определяющие состояние пальцев (сжатые или разжатые), были оптимизированы автоматически с использованием набора аннотированных данных. Оптимизация проводилась отдельно для каждой пары соседних пальцев путём максимизации точности классификации.

3.7 Графический интерфейс

Для удобства работы пользователя был разработан графический интерфейс с помощью библиотеки Tkinter на языке Python, позволяющий интерактивно загружать и обрабатывать изображения, а также настраивать параметры обработки (морфологические операции и пороги определения состояния пальцев).

4 Описание программной реализации

Программная реализация выполнена на языке программирования Python с использованием библиотек OpenCV, NumPy, SciPy, NetworkX, Scikit-image и Tkinter. Ниже подробно описаны ключевые компоненты и этапы реализации.

4.1 Общая структура реализации

Реализованное приложение представляет собой графический интерфейс, созданный с помощью библиотеки Tkinter, позволяющий выполнять сегментацию ладони, скелетизацию, распознавание кончиков пальцев и определение позы руки. Интерфейс включает элементы для загрузки изображений, настройки морфологических параметров обработки, выполнения обработки и сохранения результатов.

4.2 Чтение и предобработка изображений

Загрузка изображений выполняется функцией `read_tif_image`, которая обеспечивает совместимость с путями, содержащими русские символы, за счет предварительного чтения файла в байтовом режиме и последующей его декодировки с помощью библиотеки OpenCV.

4.3 Сегментация ладони

Сегментация реализована на основе алгоритма К-средних (K-means), выделяющего наиболее яркий кластер изображения, соответствующий ладони. Последующая морфологическая обработка включает серию операций эрозии и дилатации с регулируемыми параметрами, доступными в интерфейсе для адаптивного улучшения сегментации ладони.

4.4 Скелетизация и построение графа

На следующем этапе бинарная маска ладони подвергается скелетизации с использованием функции `skeletonize` из библиотеки Scikit-image. Полученный скелет преобразуется в графовую структуру средствами библиотеки NetworkX, где каждый пиксель скелета представляется узлом графа. Эта структура используется для поиска кончиков пальцев.

4.5 Определение кончиков пальцев

Определение кончиков пальцев реализовано следующим образом:

- На графе скелета ладони выделяются конечные узлы (с одним ребром), которые считаются кандидатами в кончики пальцев.
- Выполняется фильтрация кандидатов по угловому критерию, оставляющему только наиболее вероятные точки, соответствующие реальным пальцам.

- Кончики пальцев сортируются с использованием геометрического подхода на основе выпуклой оболочки и анализа угловых расстояний между пальцами для устойчивого определения нумерации пальцев, начиная от большого пальца.

4.6 Распознавание позы руки

Определение позы ладони осуществляется вычислением расстояний между соседними пальцами и сравнением этих расстояний с заранее определёнными пороговыми значениями, полученными в результате предварительного анализа данных. Итоговая поза кодируется последовательностью символов «+» (сомкнутые пальцы) и «-» (разомкнутые пальцы), формируя итоговый код позы формата 1*2*3*4*5, где символы «*» обозначают состояние между соседними пальцами.

4.7 Графический интерфейс

Графический интерфейс пользователя создан с помощью Tkinter и включает:

- Загрузку одного или нескольких изображений.
- Настройку параметров эрозии и дилатации через ползунки с возможностью сброса к исходным значениям.
- Просмотр исходного и обработанного изображений непосредственно в окне приложения.
- Сохранение результатов обработки (исходного изображения, обработанного изображения и текстового файла с координатами и кодом позы).

Предусмотрен также режим пакетной обработки для удобной работы с большим количеством изображений без вывода их на экран, с автоматическим сохранением результатов в выбранную пользователем директорию.

4.8 Формат сохранения результатов

Результаты сохраняются в виде изображений и текстовых файлов. Для одиночной обработки создаётся отдельный файл формата `results_<image_name>.txt`, содержащий координаты кончиков пальцев и распознанный код позы. При пакетной обработке формируется единый файл `Results.txt`, в котором для каждого обработанного изображения указываются имя файла, координаты пальцев и код позы.

5 Демонстрация работы программы

- Пользовательский интерфейс программы:

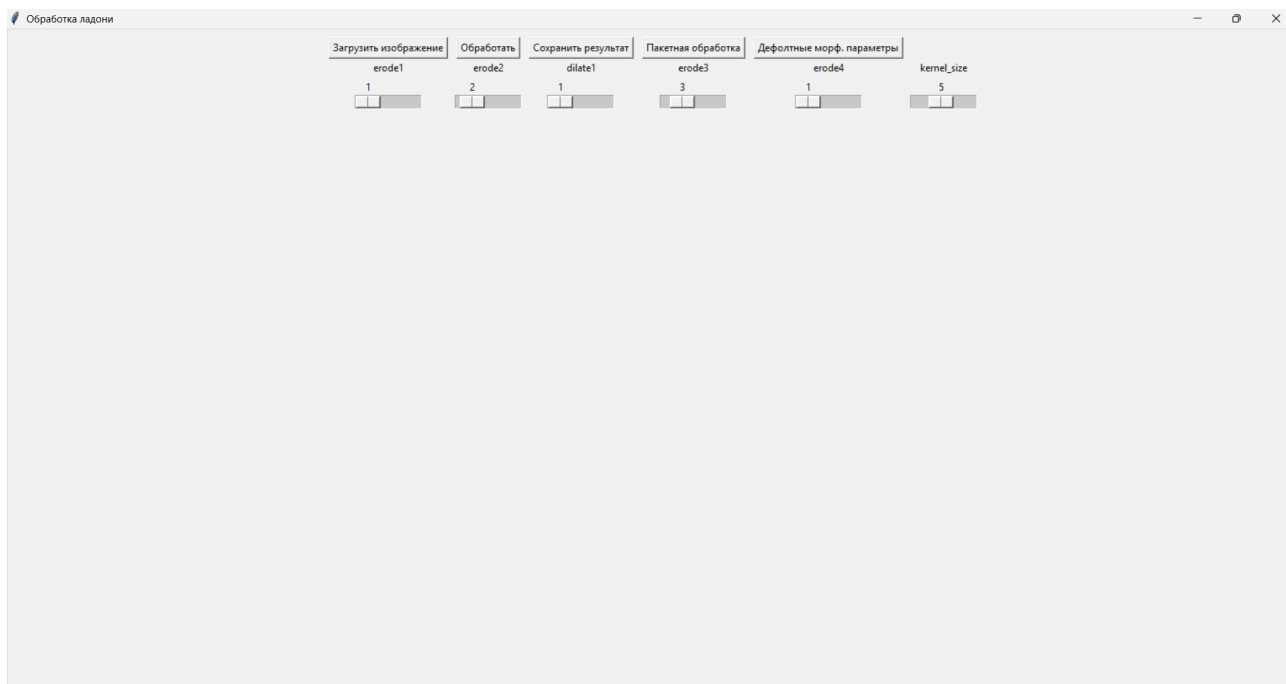


Рис. 1: Изначальный пользовательский интерфейс

- После добавления и обработки файла:

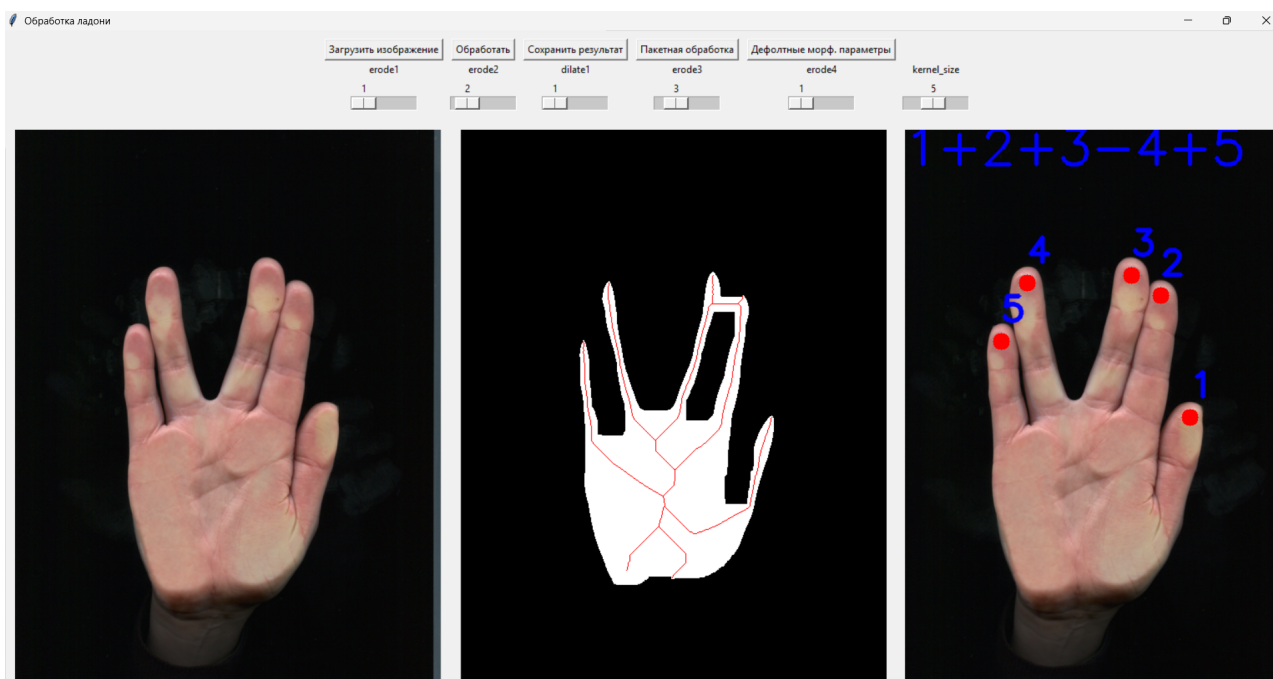


Рис. 2: Пользовательский интерфейс после выбора и обработки файла

Пример демонстрационных изображений:

- Исходное изображение:



Рис. 3: Исходное изображение, выбранное пользователем

- Маска сегментации со скелетом:

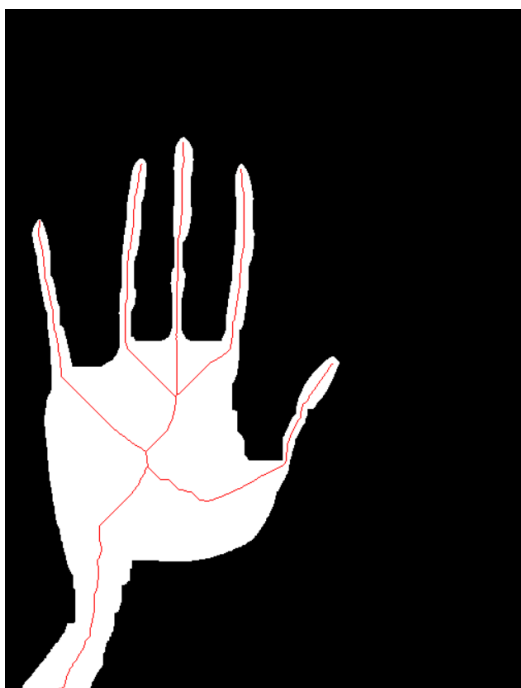


Рис. 4: Маска сегментации со скелетом

- Итоговое изображение:

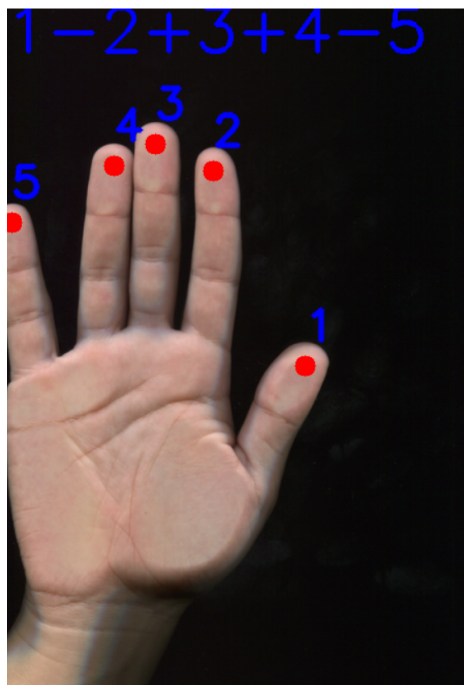


Рис. 5: Итоговое изображение с выделенными точками кончиков пальцев и анализа их расположения

6 Выводы

В результате выполнения лабораторной работы были решены задачи сегментации ладони, построения скелетного представления и автоматического распознавания позы руки на цифровых изображениях. Реализованная система выполняет поставленные задачи и предоставляет пользователю удобный интерфейс для настройки параметров обработки, позволяя адаптировать метод под особенности конкретных изображений.

Проведённый анализ результатов обработки показал, что предложенный подход на основе алгоритма К-средних и морфологических преобразований демонстрирует высокую точность сегментации ладони и устойчивое определение кончиков пальцев при подходящем выборе параметров. Алгоритм сортировки пальцев с использованием геометрических критериев и выпуклой оболочки обеспечил стабильность распознавания номеров пальцев даже в условиях шумов и неидеальной сегментации.

Дополнительно реализованный режим пакетной обработки изображений существенно упрощает работу с большими наборами данных, что является важным преимуществом при автоматизации процессов анализа изображений. Возможность гибкой настройки параметров обработки через графический интерфейс

значительно повышает удобство использования приложения.

Тем не менее, точность распознавания позы руки в отдельных случаях сильно зависит от качества входного изображения и выбранных морфологических параметров, что требует предварительной настройки и калибровки.

Список литературы

- [1] Гонсалес Р., Вудс Р., Цифровая обработка изображений.: Пер. с англ. - М.: Техносфера, 2012. - 1104 с. (3-е издание, дополненное и исправленное)