



НАЦИОНАЛЬНЫЙ ИССЛЕДОВАТЕЛЬСКИЙ
УНИВЕРСИТЕТ

Факультет компьютерных наук
Департамент программной инженерии
Выпускная квалификационная работа
«Программа имитации рукописного текста»

Научный руководитель:
доцент департамента программной
инженерии, к.т.н.
Ахметсафина Римма Закиевна

Выполнил студент группы БПИ131
образовательной программы
09.03.04 «Программная инженерия»
Семенкович Софья Алексеевна

Аффинное преобразование – отображение плоскости или пространства в себя, при котором параллельные линии остаются параллельными, пересекающиеся – пересекающимися, скрещивающиеся – скрещивающимися.

Бинаризация – приведение изображения к черно-белому виду

Глиф – конкретное изображение графемы. Одному символу может соответствовать бесконечное множество глифов

Графема – единица письменности (буква, иероглиф, цифра и т.д.)

Дерево решений – один из алгоритмов классификации, представляющий собой дерево, в листьях которого стоят значения целевой функции, а в остальных узлах – условия перехода, определяющие, по какому из ребер идти, в зависимости от того, истинно или ложно условие для классифицируемого объекта

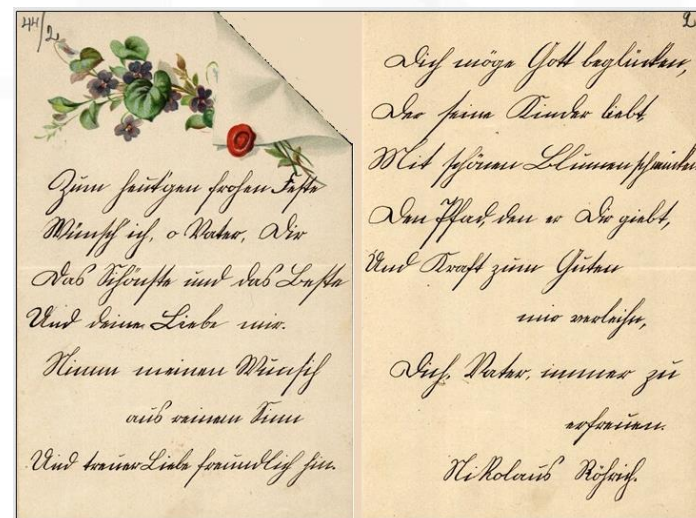
Классификация – раздел машинного обучения, посвященный решению такой задачи: существуют множество объектов, которые разделены на классы; для конечного числа объектов из этого множества известны их классы; требуется построить алгоритм, который способен определить классовую принадлежность произвольного объекта из множества. Данный алгоритм называется **классификатором**

Лигатура – линия, соединяющая два глифа. Чем больше лигатур в тексте, тем более связным он выглядит

Сегментация изображения – разделение изображения на части для облегчения последующего анализа

Собственный вектор – ненулевой вектор u , который при умножении на некоторую квадратную матрицу A превращается в самого себя с числовым коэффициентом λ , называется собственным вектором матрицы A . Коэффициент λ называется **собственным значением матрицы**

- Конспекты и рефераты
- Открытки и письма
- Игровая индустрия
- Кинематограф
- Мультипликация
- Комиксы



ОБОСНОВАНИЕ АКТУАЛЬНОСТИ РАБОТЫ

- 1977 г.: Кнут, METAFONT
- С 1980-ых г.г.: Создание множества рукописных шрифтов
- 1996 г.: Изабель Гуйон, имитация рукописного текста с использованием планшета
- 2013 г.: Алекс Грейвс, имитация рукописного текста с использованием планшета и нейронных сетей
- С 2000-ых г.г.: Создание большого количества веб-сервисов для генерации рукописного текста
- 2015 г. Веб-сервис «Handwriter»
- 2016 г.: Веб-сервис «Писец»
- 2016 г.: Томас Аода и др., “My Text in Your Handwriting”

АНАЛИЗ СУЩЕСТВУЮЩИХ РЕШЕНИЙ

Название	Программа имитации рукописного текста	Шрифты	Имитация текста с использованием планшета	Писец	Handwriter	Англо-язычные веб-сервисы
Критерий						
Поддержка латиницы	+	+	+	+	-	+
Возможность использования своего почерка	+	-	+	+	-	-
Автоматическое распознавание почерка	+	-	+	-	-	-
Возможность работы без подключения к интернету	+	+	+	-	+	-
Бесплатность	+	+/-	+	+/-	+	-
Скорость работы	высокая	высокая	высокая	высокая	средняя	высокая
Правдоподобность результатов	высокая	низкая	низкая	высокая	средняя	средняя

ЦЕЛЬ И ЗАДАЧИ РАБОТЫ

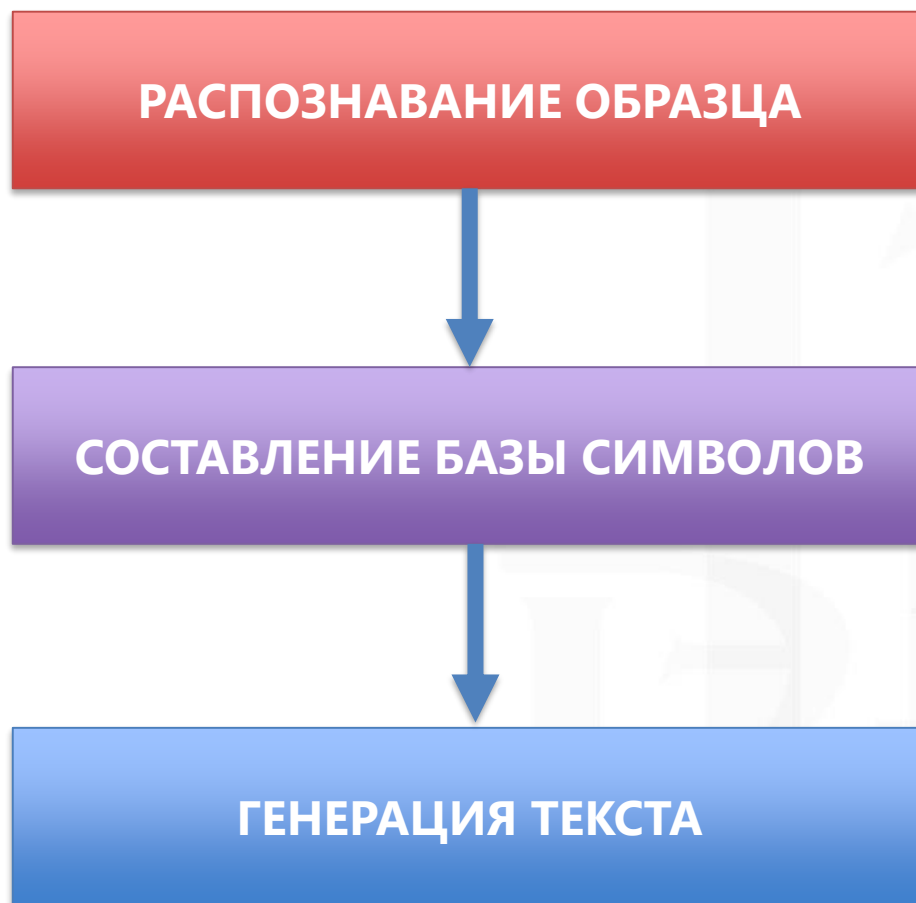
Цель работы:

Создание программного продукта, позволяющего реалистично имитировать почерк конкретного человека по предоставленному образцу.

Задачи работы:

1. Изучение предметной области.
2. Сравнение различных подходов к распознаванию изображений и синтезу почерка.
3. Выбор алгоритмов реализации, информационных моделей и структур данных.
4. Выбор технических средств реализации.
5. Разработка программного продукта, позволяющего распознавать текст и формировать банк данных из предоставленного образца (с возможностью корректировки неправильно распознанных символов), а также правдоподобно имитировать почерк.
6. Разработка технической документации.

ОСНОВНОЙ АЛГОРИТМ



АЛГОРИТМ РАСПОЗНАВАНИЯ ОБРАЗЦА: ПОДРОБНОЕ ОПИСАНИЕ

АЛГОРИТМ РАСПОЗНАВАНИЯ ОБРАЗЦА

ПОДГОТОВКА ИЗОБРАЖЕНИЯ

1. Перевод в оттенки серого
2. Шумоподавление
3. Бинаризация

СЕГМЕНТАЦИЯ ИЗОБРАЖЕНИЯ

1. **Сегментация строк**
2. **Сегментация слов**
3. **Сегментация символов**

РАСПОЗНАВАНИЕ СИМВОЛОВ

1. **Предварительная подготовка обучающей выборки**
2. **Обучение классификатора**
3. **Распознавание символов**

1. Перевод в оттенки серого:

$$I = 0.2125 * R + 0.7154 * G + 0.0721 * B$$

2. Шумоподавление:

1. Медианный фильтр
2. Фильтр Гаусса

3. Бинаризация:

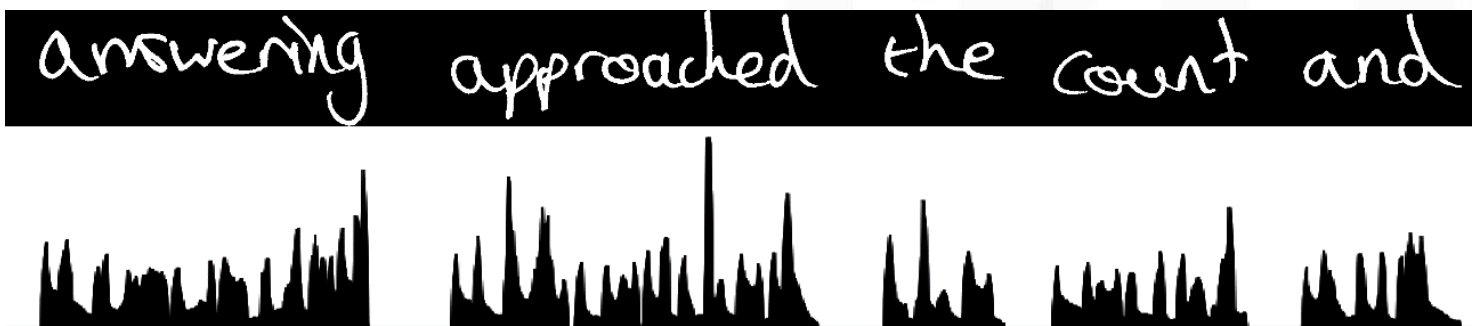
Метод Оцу

СЕГМЕНТАЦИЯ СТРОК

1. Вычислить построчные суммы значений пикселей, построить гистограмму
2. Найти области на изображении, где суммы пикселей равны нулю
 - а. Если количество таких подряд идущих строк больше определенного порога, считается, что начата новая строка.



1. Гистограмма распределения плотностей по столбцам
2. Поиск всех областей разрыва, на которых не было встречено ни одного белого пикселя в столбце
3. Количество пробелов N для данной строки в исходном тексте
4. Из всех найденных областей разрыва выбираются N максимальных, строка делится по этим областям разрыва, получившиеся области изображения считаются словами



Найти центральные моменты изображения второго порядка:

$$\mu_{02} = \frac{m_{02}}{m_{00}} - \bar{y}^2$$
$$\mu_{11} = \frac{m_{11}}{m_{00}} - \bar{x}\bar{y}$$

$$\bar{x} = \frac{m_{10}}{m_{00}}, \bar{y} = \frac{m_{01}}{m_{00}},$$

m_{00} – площадь, занятая ненулевыми (белыми) пикселями,
 m_{10} – среднее значение x по всему объекту,
 m_{01} – среднее значение y

Угол перекоса:

$$\alpha = \frac{\mu_{11}}{\mu_{02}}$$

Матрица преобразования:

$$\begin{bmatrix} 1 & 0 \\ \alpha & 1 \end{bmatrix}$$

Новые координаты точек изображения:

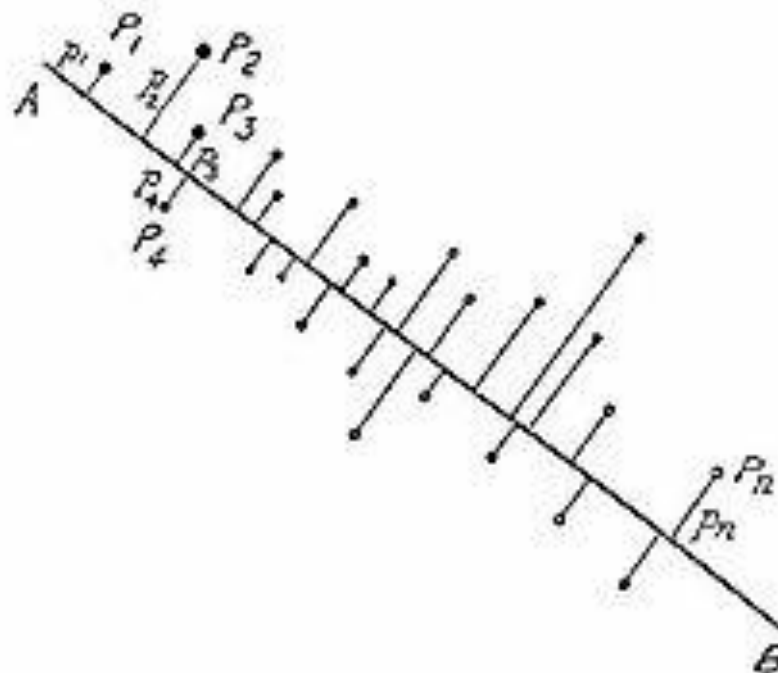
$$(x', y') = \begin{bmatrix} 1 & 0 \\ \alpha & 1 \end{bmatrix} * (x, y)$$

ПОДГОТОВКА ТРЕНИРОВОЧНОЙ ВЫБОРКИ

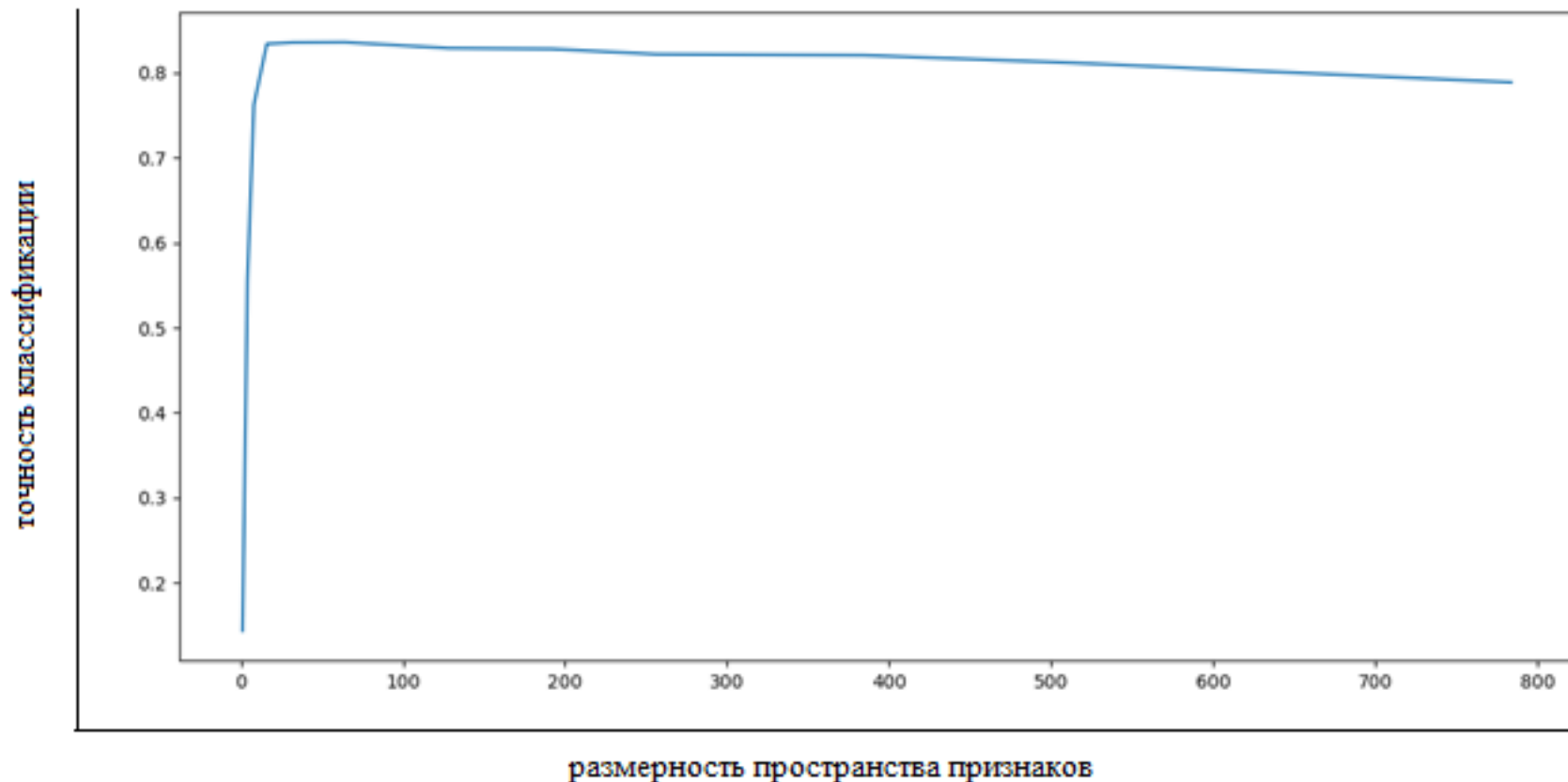
1. База данных EMNIST (2017, латинские рукописные буквы и цифры)
2. Шаги подготовки:
 - a. Бинаризация
 - b. Устранение перекоса
 - c. Уменьшение размерности пространства признаков: метод анализа главных компонент

ПОДГОТОВКА ТРЕНИРОВОЧНОЙ ВЫБОРКИ: МЕТОД АНАЛИЗА ГЛАВНЫХ КОМПОНЕНТ

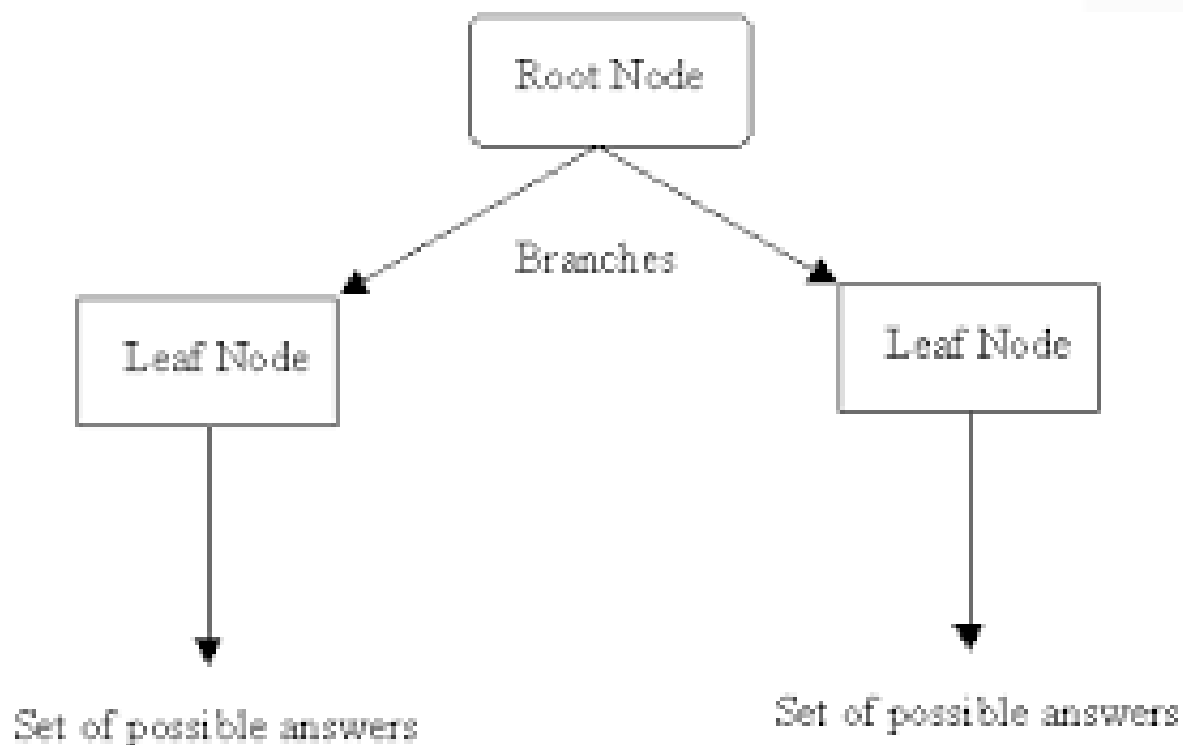
1. Кластерный анализ данных
2. По каким признакам различаются объекты?
3. Как взаимосвязаны признаки?



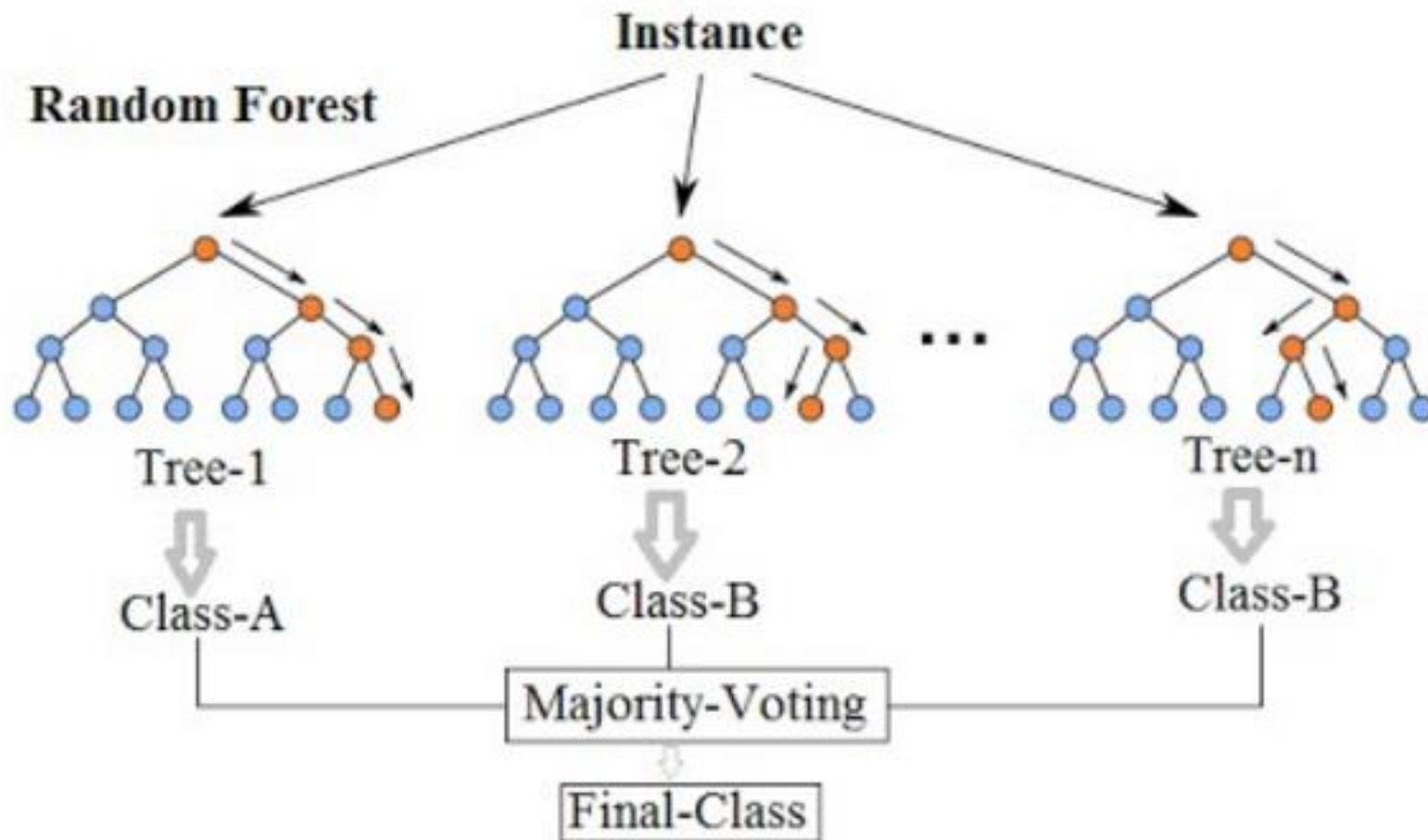
ПОДГОТОВКА ТРЕНИРОВОЧНОЙ ВЫБОРКИ: МЕТОД АНАЛИЗА ГЛАВНЫХ КОМПОНЕНТ



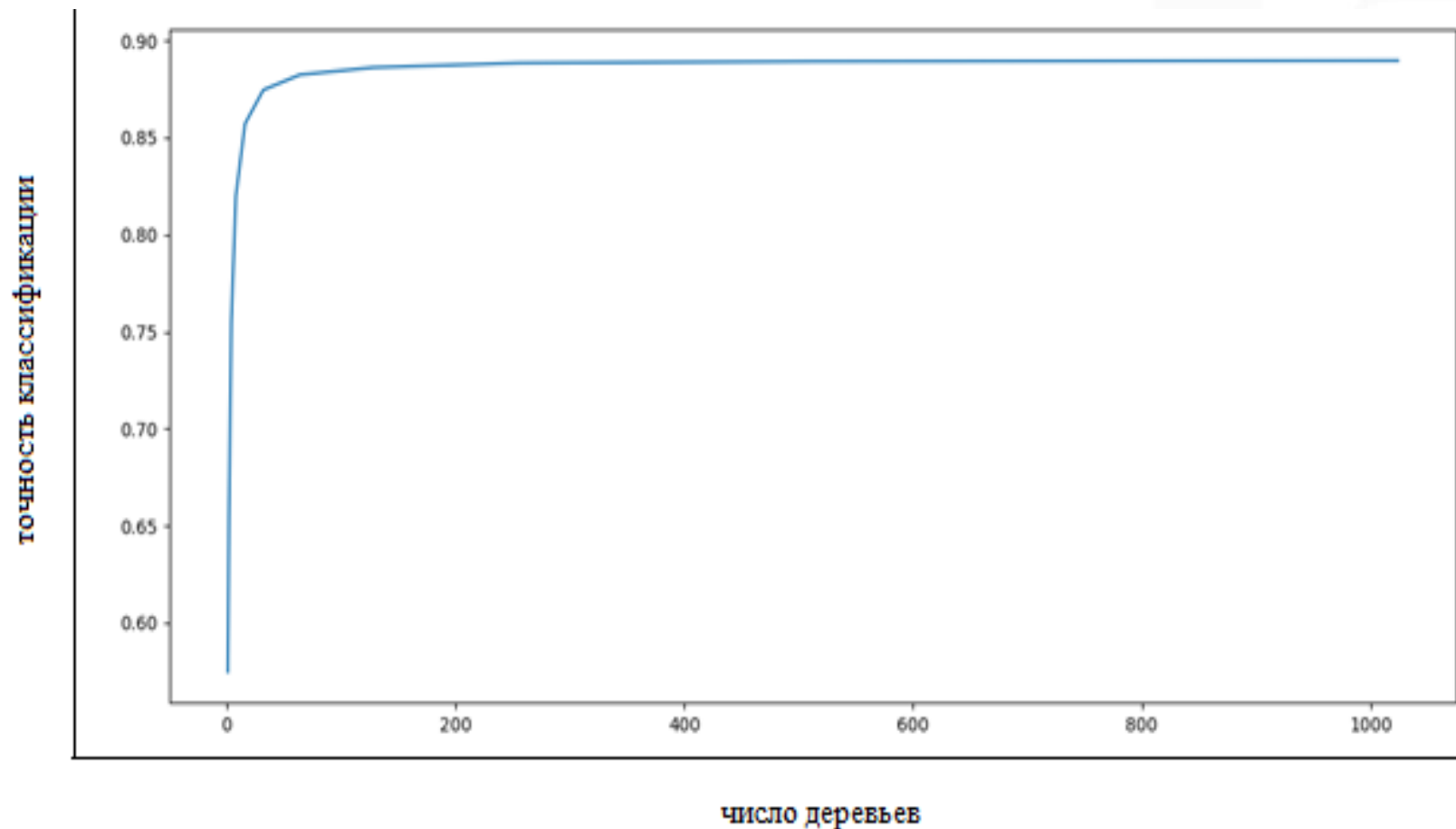
ОБУЧЕНИЕ КЛАССИФИКАТОРА: РЕШАЮЩЕЕ ДЕРЕВО



ОБУЧЕНИЕ КЛАССИФИКАТОРА: СЛУЧАЙНЫЙ ЛЕС



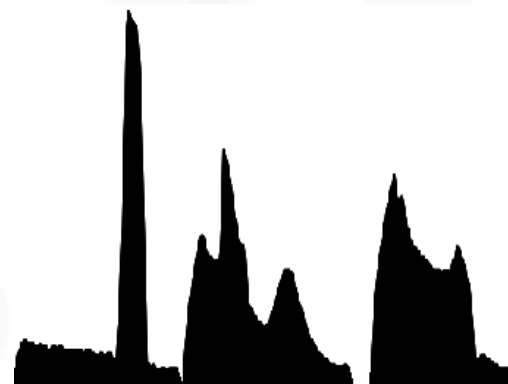
ОБУЧЕНИЕ КЛАССИФИКАТОРА: ЭКСПЕРИМЕНТ



СЕГМЕНТАЦИЯ СИМВОЛОВ

Разделение в областях, которые:

1. имеют значение плотности пикселей меньше некоторого порогового значения (0.25 от общей высоты строки);
2. расположены в предполагаемых местах разбиения символов
3. предшествуют возрастанию плотности пикселей текста



1. Вычисляется степень уверенности классификатора в том, что данное изображение является тем символом, который программа пытается отделить от слова на данном шаге
2. Результат распознавания запоминается, и программа переходит к распознаванию следующего символа
3. Разбиение с наилучшим результатом считается правильным. Такое разбиение ищется методом Дейкстры в графе узлами которого будут значения координаты по оси x предполагаемых точек разбиения слова, а веса ребер будут равны вероятности ошибки классификации данного символа
4. Для выявления отделенных от слова символов на изображении используется метод прослеживания контура



АЛГОРИТМ СОЗДАНИЯ БАЗЫ СИМВОЛОВ: ПОДРОБНОЕ ОПИСАНИЕ

СОСТАВЛЕНИЕ БАЗЫ СИМВОЛОВ

Символы в базе – это файлы, вырезанные из исходного изображения и сохраненные в соответствующие папки:

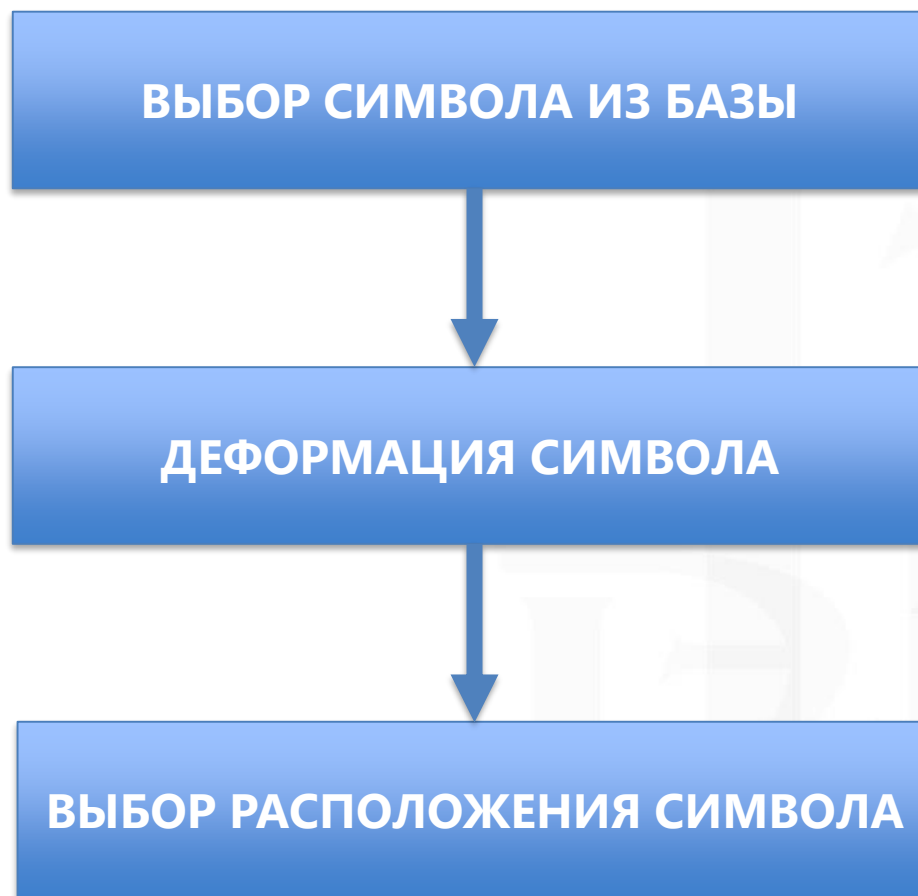
- строчные символы – в папку “lowercase\SYMBOL_NAME”
- заглавные символы – в “uppercase\SYMBOL_NAME”
- числа – в “numbers\SYMBOL_NAME”

Имена файлов имеют вид «POSITION_NEIGHBOURS_NUMBER.png», где:

- POSITION может принимать значения [“o”, “s”, “m”, “e”] (что означает: единственный символ (без соседей), в начале слова, в середине слова, в конце слова, соответственно)
- NEIGHBOURS – предыдущий и последующий соседи символа (принимает значение 0, если соседей нет)
- NUMBER – номер символа в базе

АЛГОРИТМ ГЕНЕРАЦИИ ТЕКСТА: ПОДРОБНОЕ ОПИСАНИЕ

АЛГОРИТМ ГЕНЕРАЦИИ



АЛГОРИТМ ГЕНЕРАЦИИ: ВЫБОР СИМВОЛА ИЗ БАЗЫ

При выборе символа из базы учитывается:

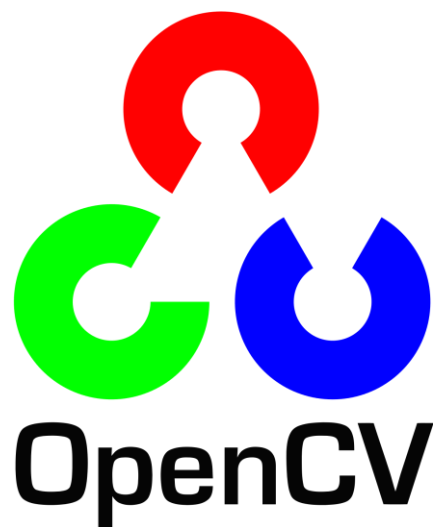
1. Класс символа
2. Положение символа в слове
3. Соседи символа

АЛГОРИТМ ГЕНЕРАЦИИ: ДЕФОРМАЦИЯ И РАСПОЛОЖЕНИЕ СИМВОЛА



ТРЕБОВАНИЯ К ПРОГРАММЕ

1. программа должна уметь распознавать текст на английском языке, написанный от руки (английские большие и маленькие буквы, а также цифры)
 - a. инструмент написания – шариковая или капиллярная ручка
 - b. текст должен быть написан на обычном белом листе бумаги формата А4, отсканированном с разрешением 600 dpi
 - c. строки на изображении не должны пересекаться между собой
 - d. на вход программе должен подаваться также файл с транскрипцией текста на изображении
2. программа должна составлять базу символов из образца почерка (описание базы представлено ниже)
3. программа должна генерировать текст, который будет предоставлен пользователем, почерком образца:
 - a. текст должен состоять только из тех символов, которые присутствовали в образце
 - b. программа должна генерировать .png-файл с готовым текстом



ДЕМОНСТРАЦИЯ

ОСНОВНЫЕ РЕЗУЛЬТАТЫ РАБОТЫ

1. Изучены источники информации по компьютерному зрению, обработке изображений и машинному обучению, получено достаточно глубокое представление об этих областях;
2. Рассмотрены и проанализированы различные подходы к распознаванию и синтезу рукописного текста;
3. Выбраны наиболее подходящие для использования при реализации данного приложения алгоритмы
4. Выбраны технические средства реализации
5. Разработано приложение, которое:
 - a. Распознает предоставленный пользователем образец почерка,
 - b. Позволяет корректировать неправильно распознанные символы,
 - c. Составляет на его основе базу символов,
 - d. Генерирует запрашиваемый пользователем текст с использованием данной базы.
6. Разработана техническая документация

НАПРАВЛЕНИЯ ДАЛЬНЕЙШЕЙ РАБОТЫ

1. Реализация поддержки кириллицы
2. Добавление знаков пунктуации
3. Обеспечение распознавания символов без предоставления пользователем транскрипта рукописного текста

СПИСОК ИСПОЛЬЗОВАННЫХ ИСТОЧНИКОВ

1. Дёмин А. Ю., Кудинов А. В. Цифровые фильтры изображений [Электронный ресурс] / Компьютерная графика. Режим доступа: <http://compgraph.tpu.ru/filters.htm>, свободный. (дата обращения 25.04.17)
2. Drizhuk D. Computer Graphics [Электронный ресурс] / complynx.net. Режим доступа: <http://complynx.net/4/ip.pdf>, свободный (дата обращения 20.05.17)
3. Cohen G., Afshar S., Tapson J., Schaik A. EMNIST: an extension of MNIST to handwritten letters [Электронный ресурс]: arXiv, 2017. – Режим доступа: <https://arxiv.org/abs/1702.05373>, свободный. (дата обращения: 20.04.17)
4. Haines T.S.F., Aodha O.M., Brostow G. J. My Text in Your Handwriting/ ACM Transaction on Graphics - 2016. - Vol. 35.
5. Турбин А. Программная реализация размытия по Гауссу // РГРТУ - 2007.
6. Демин А.А. Обзор интеллектуальных систем для оценки каллиграфии // Москва : Инженерный вестник (МГТУ им. Н.Э. Баумана). Электронный журнал. - 2012 - № 9.
7. Исследование методов обработки изображений [Электронный ресурс] / Вунивере.ру. Режим доступа: <https://vunivere.ru/work15671/page6>, свободный (дата обращения: 20.04.17).
8. Р.К. Захаров Методы повышения качества изображений в задачах распознавания // Современные научные исследования и инновации. - 2012 - 8.



НАЦИОНАЛЬНЫЙ ИССЛЕДОВАТЕЛЬСКИЙ
УНИВЕРСИТЕТ

Спасибо за внимание!

Семенкович Софья Алексеевна,
sasemenkovich@yandex.ru

Москва - 2017

ПЕРЕВОД В ОТТЕНКИ СЕРОГО

- Изображение - матрица пикселей;
- Каждый пиксель обладает степень интенсивности красного, зеленого и синего цветов по шкале от 0 до 255;
- Перевод в оттенки серого осуществляется по формуле:

$$I = 0.2125 * R + 0.7154 * G + 0.0721 * B$$

ПЕРЕВОД В ОТТЕНКИ СЕРОГО

The two gentlemen haistily seperated
each in a fever of excitement wholly



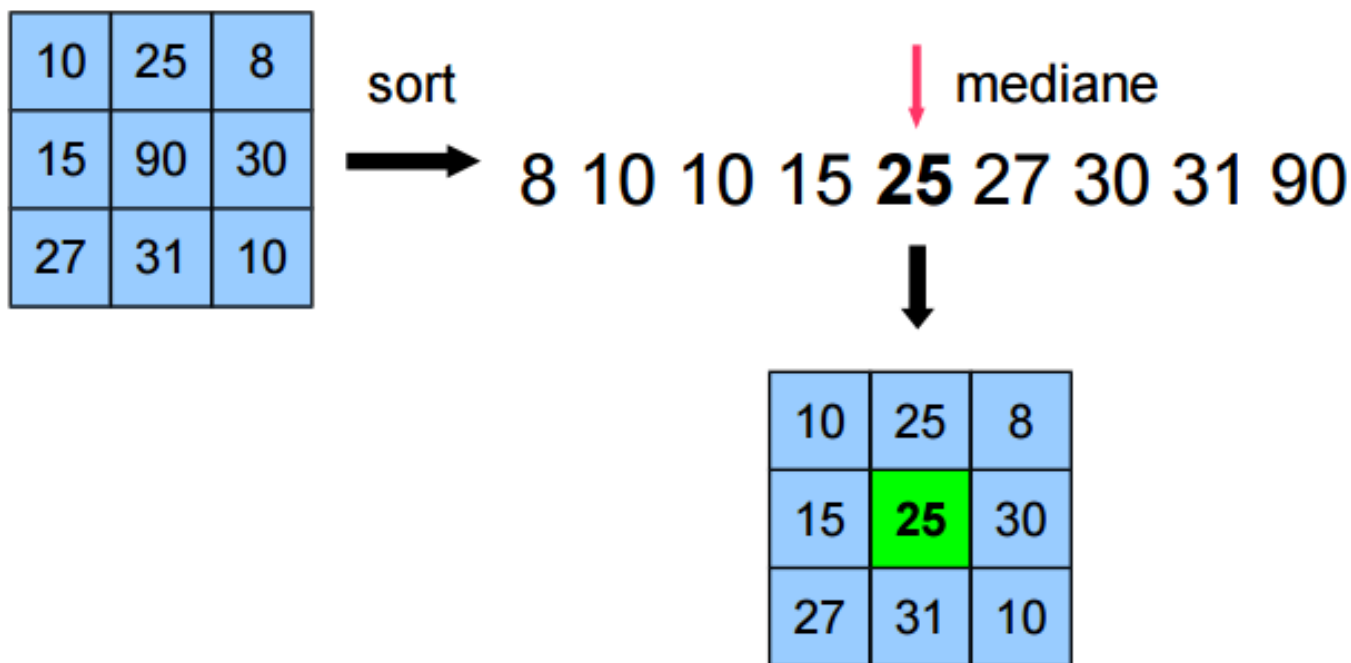
The two gentlemen haistily seperated
each in a fever of excitement wholly

ШУМОПОДАВЛЕНИЕ: ВИДЫ ШУМОВ

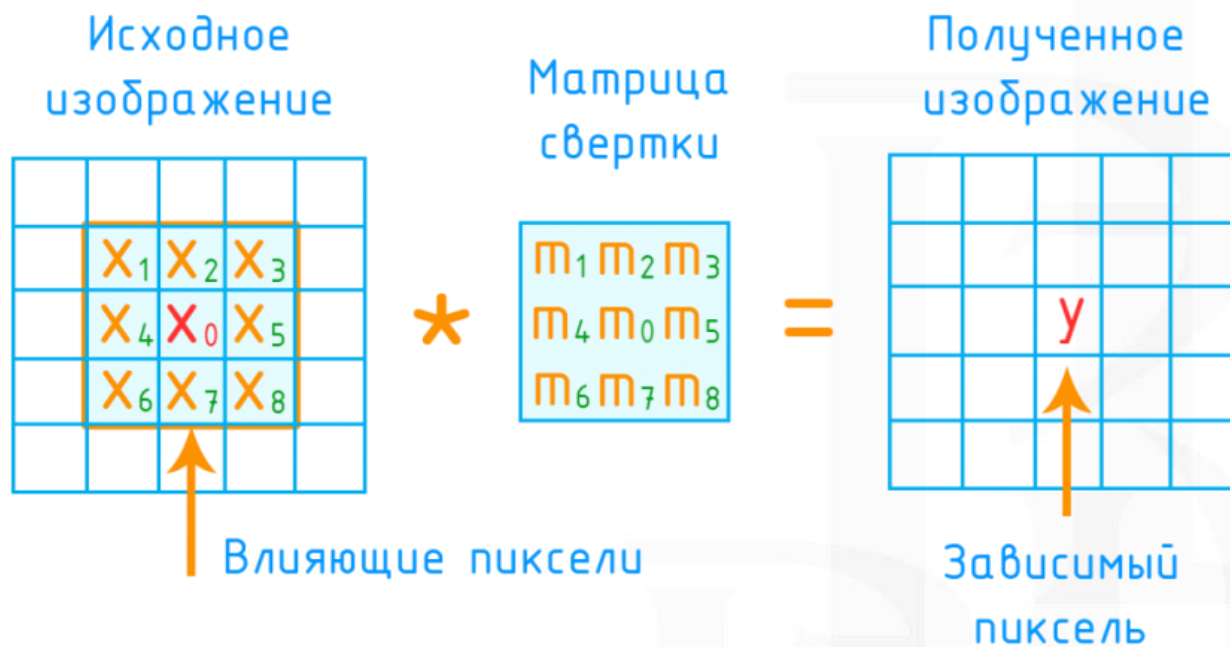
- Импульсные (соль и перец)
- Аддитивные (гауссов шум)
- Мультипликативные



ШУМОПОДАВЛЕНИЕ: МЕДИАННЫЙ ФИЛЬТР



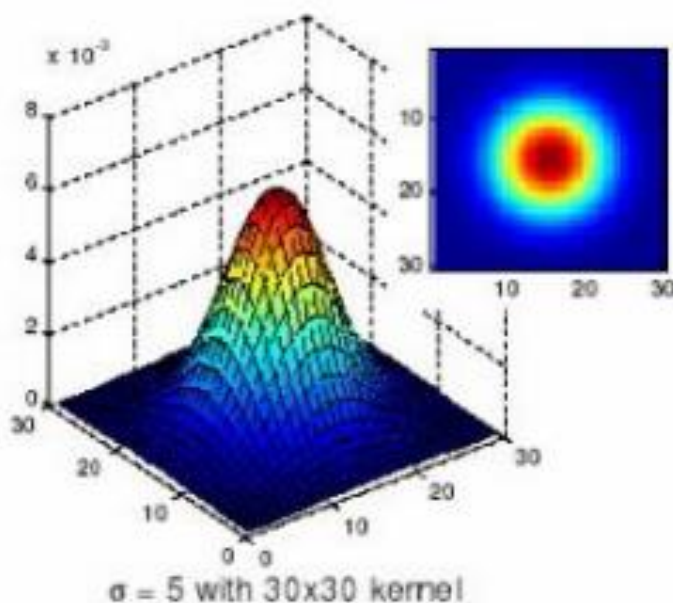
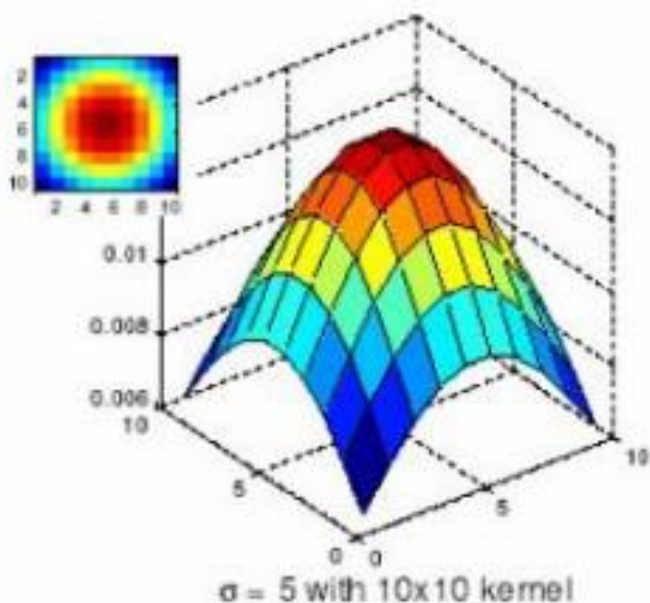
ШУМОПОДАВЛЕНИЕ: ФИЛЬТР ГАУССА



ШУМОПОДАВЛЕНИЕ: ФИЛЬТР ГАУССА

$$G(u, v) = \frac{1}{2\pi\sigma^2} e^{-\frac{(u^2+v^2)}{2\sigma^2}}$$

0,003	0,013	0,022	0,013	0,003
0,013	0,059	0,097	0,059	0,013
0,022	0,097	0,159	0,097	0,022
0,013	0,059	0,097	0,059	0,013
0,003	0,13	0,022	0,013	0,003



БИНАРИЗАЦИЯ: МЕТОД ОЦУ

