

# 1 Аннотация

В этой работе были исследованы классические методы автоматической бинаризации изображений. Была написана программа, демонстрирующая работу популярных методов глобальной и адаптивной бинаризации. Также был сделан веб-сайт, чтобы каждый мог самостоятельно бинаризовать изображения в пару кликов.

## 2 Описание различных методов бинаризации

Главной целью бинаризации является радикальное уменьшение количества информации, с которой приходится работать. Удачная бинаризация сильно упрощает последующую работу с изображением.



Существуют различные методы бинаризации, которые можно условно разделить на две группы:

- глобальные (пороговые);
- локальные (адаптивные).

### 2.1 Глобальные методы бинаризации

В глобальных методах бинаризации происходит работа со всем изображением сразу. В ходе работы находится порог бинаризации  $t$ , с помощью которого происходит деление на черное и белое, причем величина порога  $t$  остается неизменной в течение всего процесса бинаризации.

#### Метод Оцу

Популярным методом глобальной бинаризации изображений является метод Оцу.

С помощью данного метода вычисляется порог  $t$ , минимизирующий среднюю ошибку сегментации, т.е. среднюю ошибку от принятия решения о принадлежности пикселей изображения объекту или фону. Значения яркостей пикселей изображения можно рассматривать как случайные величины, а их гистограмму – как оценку плотности распределения вероятностей. Если плотности распределения вероятностей известны, то можно определить оптимальный (в смысле минимума ошибки) порог для сегментации изображения на два класса  $c_0$  и  $c_1$  (объекты и фон).

Гистограмма строится по значениям  $p_i = \frac{ni}{N}$ . В данной формуле  $N$  – общее количество пикселей изображения с уровнем яркости  $i$ . Порог  $t$  представляет собой целое значение от 0 до  $L=\max$ . При помощи гистограммы все пиксели разделяются на «полезные» (объектные) и фоновые. Каждому виду соответствуют относительные частоты  $W_0$  и  $W_1$ :

$$W_0(t) = \sum_{i=1}^t p_i \quad (1)$$

$$W_1(t) = \sum_{i=t+1}^L p_i = 1 - W_0(t) \quad (2)$$

Далее вычисляются средние уровни для каждого вида изображения по формулам:

$$\mu_0(t) = \sum_{i=1}^t \frac{ip_i}{W_0} \quad (3)$$

$$\mu_1(t) = \sum_{i=t+1}^L \frac{ip_i}{W_1} \quad (4)$$

Далее ищется порог, который уменьшает дисперсию внутри вида пикселей, определяемую следующей формулой:

$$\delta_W^2(t) = W_1(t)\delta_1^2(t) + W_2(t)\delta_2^2(t) \quad (5)$$

Следующим шагом определяется межклассовая дисперсия, по формуле, представленной ниже:

$$\sigma_{\text{кл}}^2(t) = W_0(t)W_1(t) * (\mu_1(t) - \mu_0(t))^2 \quad (6)$$

Затем вычисляется максимальное значение для оценки качества деления изображения на две части, которое соответствует искомому порогу:

$$\eta(t) = \max \left[ \frac{\sigma_{\text{кл}}^2(t)}{\delta_W^2(t)} \right] \quad (7)$$

Достоинствами метода Оцу являются:

- простота реализации;
- адаптация к различного рода изображениям, при помощи выбора оптимального порога;
- быстрое время выполнения.

## ISODATA

Алгоритм ISODATA является разновидностью итерационных алгоритмов, в основе которого лежит метод разделения-объединения кластеров.

Алгоритм формирования кластеров строится следующим образом. Имеется множество кластеров  $C_1, C_2, \dots, C_k$  с математическими ожиданиями  $m_1, m_2, \dots, m_k$  и ковариационная матрица  $\sum_k$  для  $k$ -го кластера.

Шаги алгоритма:

1. Поместить вектор  $x_i$  в кластер  $l$ , для которого достигается минимум

$$D_M = (x_i - m_l)' \sum_l^{-1} (x_l - m_l)$$

2. Объединить кластеры  $i$  и  $j$ , если

$$|m_i - m_j| < \tau_\nu$$

где  $\tau_\nu$  – пороговое значение дисперсии.

3. Разделить кластер  $k$ , если максимальное собственное значение  $\sum_k$  превышает порог  $\tau_\nu$ .

4. Завершить работу, если для каждого кластера  $I$  выполняется условие

$$|m_i(t) - m_i(t+1)| < \varepsilon$$

или если достигнуто максимально допустимое значение итераций.

## 2.2 Локальные методы бинаризации

Локальные (адаптивные) методы бинаризации производят разбиение изображения на несколько областей, для каждой из которых необходимо вычислить порог, основываясь на информации об интенсивности пикселей.

Алгоритмы данного класса предполагают разбиение изображения на блоки определенного размера, при этом размер блока должен быть минимальным, но достаточным для сохранения исходных особенностей и деталей изображения. Однако при этом блоки должны быть настолько большими, чтобы шумы влияли на результат минимально. Функция сглаживания результирующего растра при адаптивной бинаризации позволяет получить удовлетворительный результат без использования дополнительных фильтров.

### Метод Ниблэка

В методе Ниблэка для каждого пикселя изображения необходимо получить свое значение порога. Его величина определяется на основе вычисления локального среднего и локального среднеквадратического отклонения. Значение порога для точки с координатами  $(m, n)$  считается по следующей формуле:

$$t(m, n) = \mu(m, n) + k * \sigma(m, n) \quad (8)$$

В данной формуле  $\mu(m, n)$  представляет собой среднее,  $\sigma(m, n)$  – среднеквадратичное отклонение в локальной окрестности рассматриваемой точки изображения, а значение  $k$  определяет, какую именно часть границы объекта необходимо взять в качестве объекта. Метод Ниблэка за счет своей простоты позволяет достичь наиболее высокую скорость бинаризации изображений. Метод используется на практике для быстрой фильтрации контрастных изображений, на которых практически отсутствуют сильно зашумленные участки с плавными переходами яркости.

### Метод Саувола

К методам локальной адаптивной бинаризации относится и метод Саувола. Определение локального порога бинаризации осуществляется с помощью прохождения всего изображения окном  $w \times w$ . В методе бинаризации Саувола порог  $t(x, y)$  определяется следующей формулой, используя для вычисления среднее значение  $m(x, y)$  и среднеквадратическое отклонение  $s(x, y)$  интенсивности пиксела в окне  $w \times w$  вокруг пиксела  $(x, y)$ :

$$t(x, y) = m(x, y) \left[ 1 + k \left( \frac{s(x, y)}{R} - 1 \right) \right] \quad (9)$$

В данной формуле  $R$  представляет максимальное отклонение ( $R = 128$  для изображения в оттенках серого), а  $k$  является параметром, который принимает значения в диапазоне  $[0.2, 0.5]$ .

Метод Саувола широко применяется к изображениям, в которых яркость изображения распределяется неравномерно. Однако, алгоритм Саувола менее устойчив к зашумленности исходного изображения, чем, к примеру, алгоритм Оцу. Также у метода имеются трудности с изображениями, у которых мало освещения, особенно в случаях, когда значения пикселей объекта находятся близко друг к другу. При обработке тонких пересекающихся линий могут возникать разрывы, поэтому метод хорош для толстых линий и крупных объектов.

В общем случае адаптивную бинаризацию можно рекомендовать в случае, если необходимо обработать для обработки полутоновых изображений невысокого качества (сканированных снимок), на которых из-за неравномерности фона обычная бинаризация дает плохие результаты.

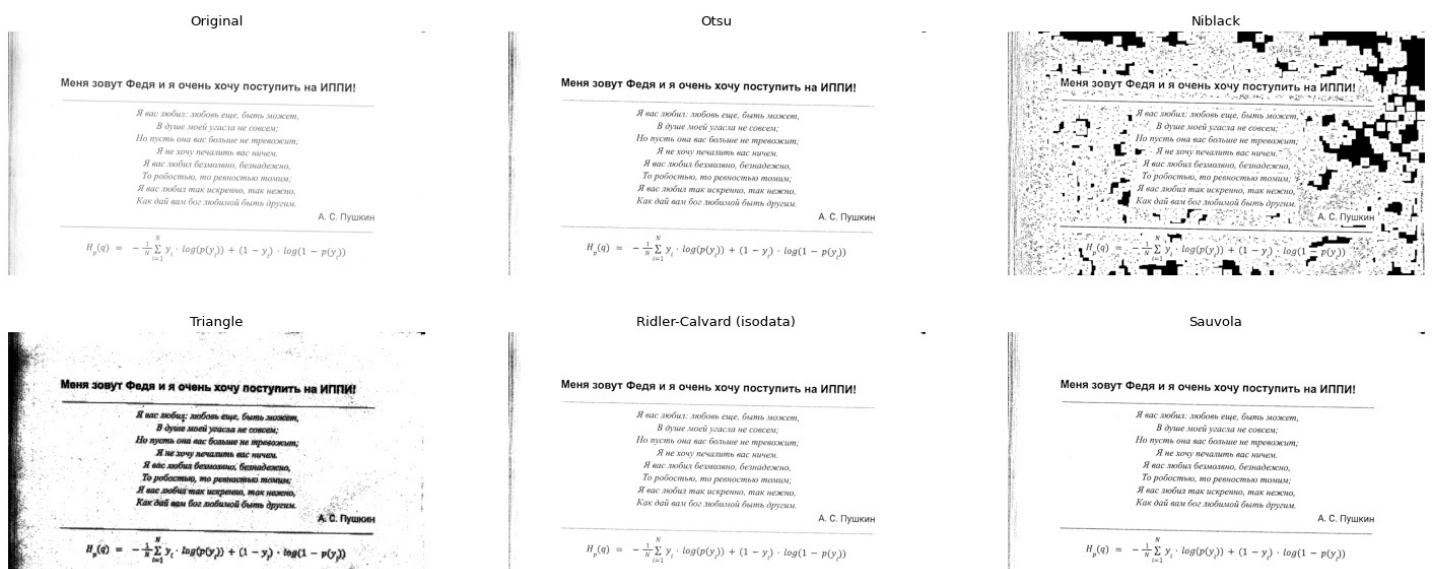
### 3 Реализация программы для демонстрации методов

Для демонстрации работы написанной программы, мы собрали DataSet из разнообразных фотографий и пропустили их через наш код. Полученные результаты исполнения можно наблюдать ниже.

#### Скан текста

Для начала был набран, распечатан и отсканирован текст. Его скан мы подали реализованным методам нашей программы. Как можно видеть из изображения ниже, все методы, за исключением Niblack, хорошо справились с поставленной задачей бинаризации скана распечатанного текста. Полученный результат можно объяснить ограниченной областью применения метода Niblack.

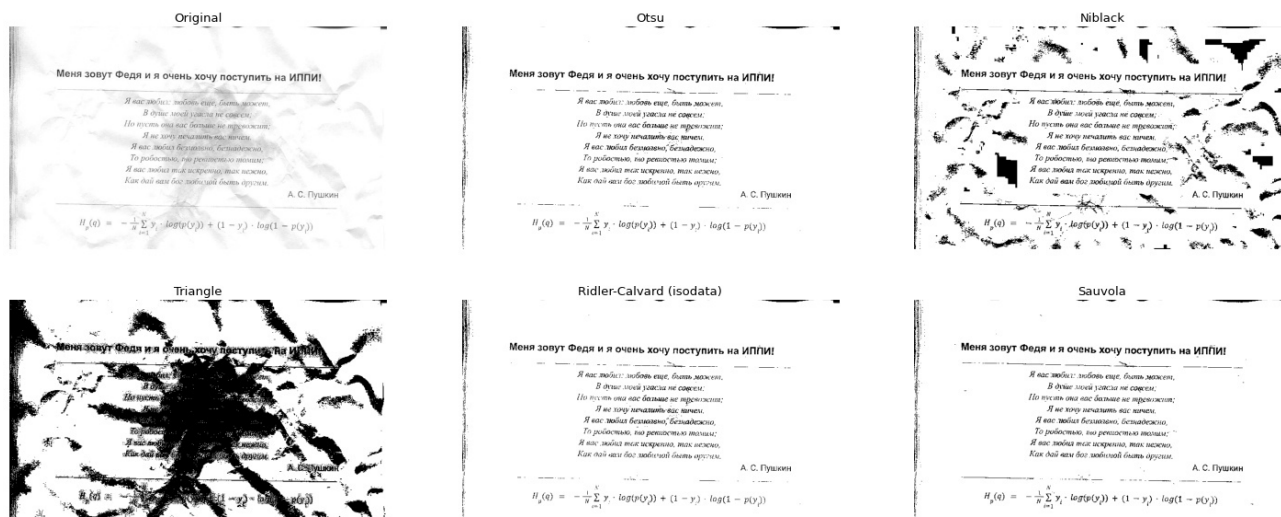
#### Скан текста



## Скан текста с дефектами

Далее нам стало интересно, как данные методы бинаризации изображения справляются с различными дефектами на фотографии. Для этого мы разлили чай на листок с текстом и скомкали его, а потом отсканировали ещё раз. Как можно наблюдать из результатов применения методов, адаптивные справились с бинаризацией хуже из-за шумов. Так же можно заметить, что дефекты на скане практически никак не отобразились на работе глобальных методов.

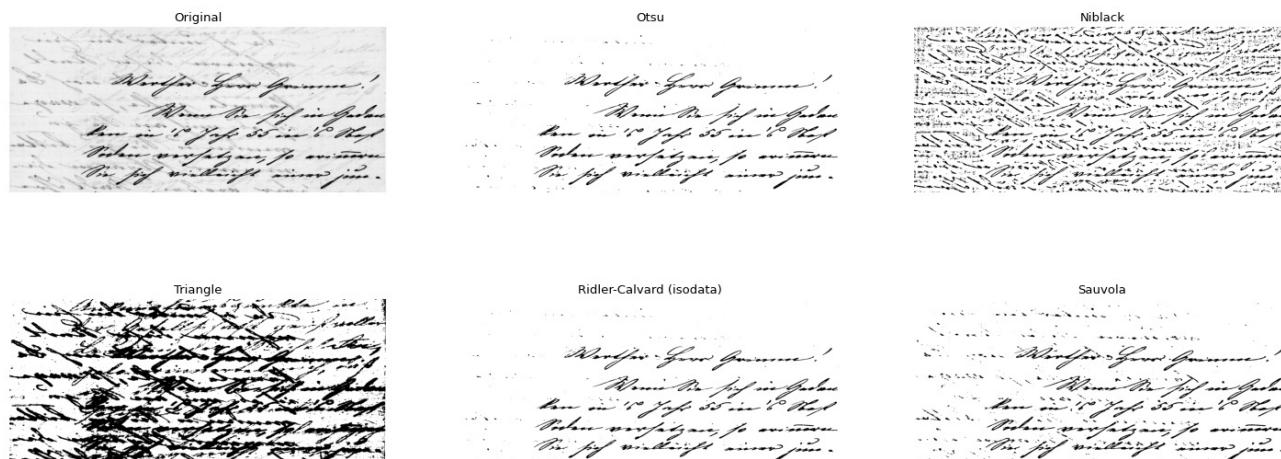
### Скан текста с дефектами



## DIBCO 2016

DIBCO - самый популярный конкурс по бинаризации рукописного и печатного текста. Мы решили посмотреть, как данные методы справятся с фотографией, предложенной на одном из соревнований. Основная сложность оригинального изображения заключается в просвечивании текста с другой стороны листа. Методы Оцу и Isodata отлично справились с этим. А вот результаты применения адаптивных методов, особенно Niblack, имеют на своём фоне большое количество шумов, мешающее нормальному распознаванию текста.

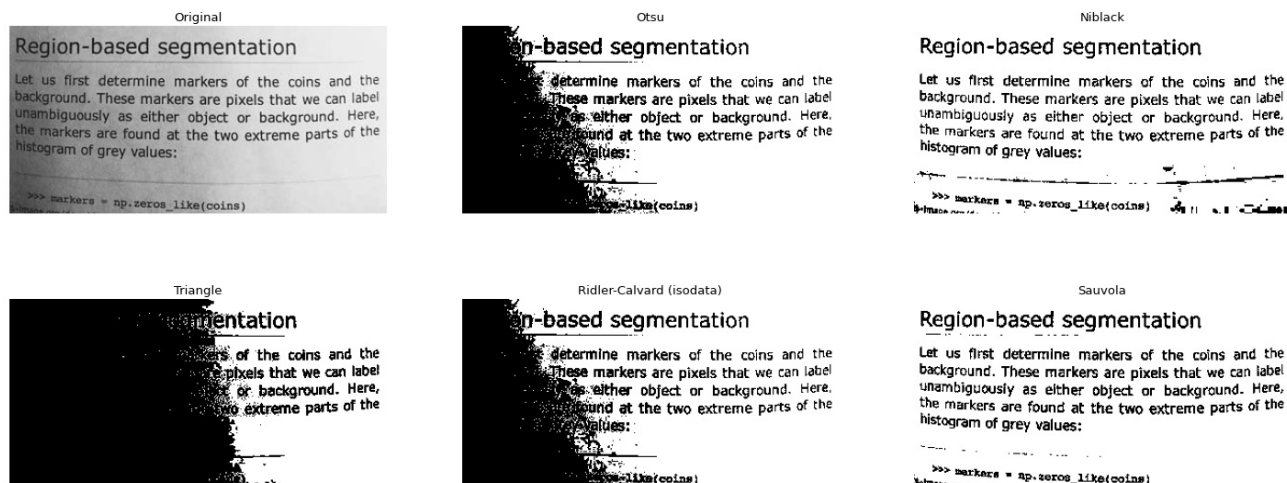
### DIBCO 2016



## Текст с неоднородной яркостью фона

Слабым местом пороговых методов является неоднородная яркость фона. Это можно наблюдать на изображении ниже. В области тени, после применения методов Оцу и Isodata, возникает тёмное пятно. Адаптивные же методы хорошо справляются с этой задачей, так как они работают с окнами.

### Текст с неоднородной яркостью фона

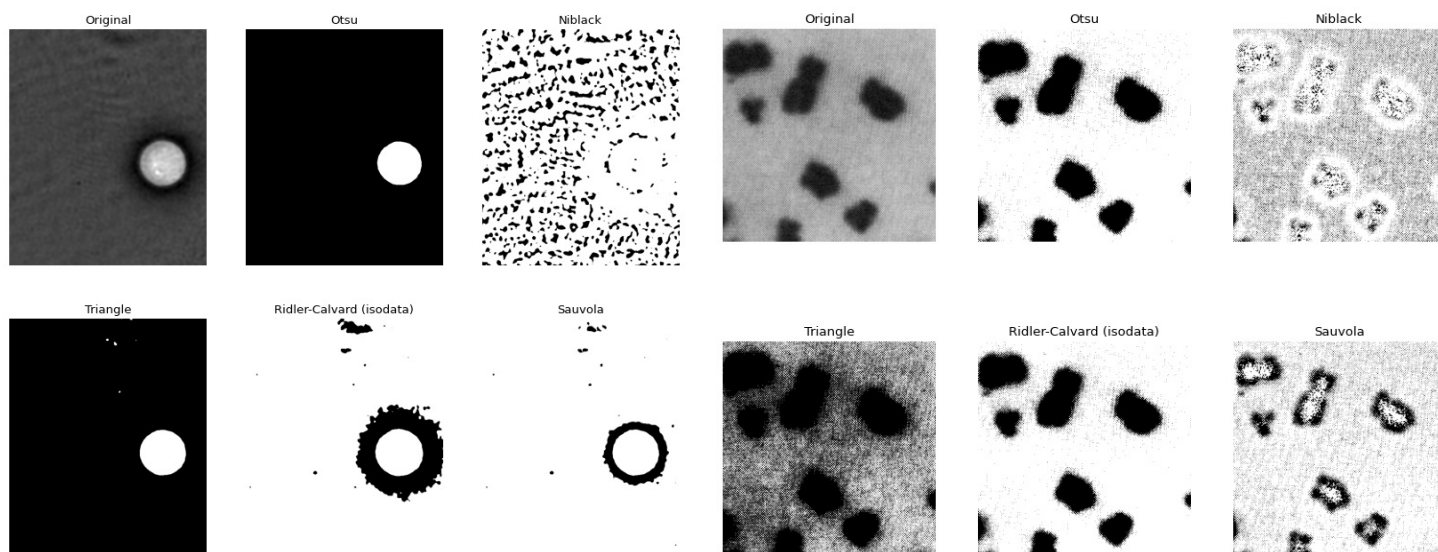


## Медицинские изображения

Также нам было интересно посмотреть, как хорошо данные методы справляются с бинаризацией медицинских изображений. Так как на таких фотографиях обычно большое количество шумов, то для них желательно использовать пороговые методы. А лучший результат даёт метод Оцу. Изображения ниже, с применением данных методов для медицинских фотографий, подтверждают выше сказанное.

### Хромосомы делятся митозом

#### Клетка человека



## 4 Вывод

Неудачи в процессе бинаризации могут привести к искажениям, таким, как разрывы в линиях, потеря значащих деталей, нарушение целостности объектов, появление шума и непредсказуемое искажение символов из-за неоднородностей фона.

Различные методы бинаризации имеют свои слабые места: так, например, метод Оцу может приводить к утрате мелких деталей и «слипанию» близлежащих символов, а метод Ниблэка грешит появлением ложных объектов в случае неоднородностей фона с низкой контрастностью. Из этого можно сделать вывод, что каждый из рассмотренных методов должен быть применен в определенной области, а также для каждой области могут быть разработаны более подходящие методы.

## Список литературы/References

1. *Исрафилов Х.С.* Исследование методов бинаризации изображений.
2. *N. Otsu.* A threshold selection method from gray-level histograms (англ.) // IEEE Trans. Sys., Man., Cyber. : journal. — 1979. — Vol. 9. — P. 62—66
3. *Селянкин В.В, Скорород С.В.* Анализ и обработка изображений в задачах компьютерного зрения - с.65-66
4. Сегментация изображения. [Электронный ресурс] <https://habr.com/ru/post/128768/>