

УДК 004.021

## СРАВНИТЕЛЬНЫЙ АНАЛИЗ МЕТОДОВ БИНАРИЗАЦИИ ИЗОБРАЖЕНИЙ



**Д.В. Орлов**  
Студент БГУИР



**С.Н. Нестеренков**  
Кандидат технических наук,  
доцент, декан факультета  
компьютерных систем и сетей



**А. Н. Марков**  
Старший преподаватель  
магистр технических наук,  
заместитель начальника  
Центра информатизации и  
инновационных разработок  
БГУИР

Центр информатизации и инновационных разработок Белорусского государственного университета информатики и радиоэлектроники, Республика Беларусь  
E-mail: [dimm4work@gmail.com](mailto:dimm4work@gmail.com), [s.nesterenkov@bsuir.by](mailto:s.nesterenkov@bsuir.by), [a.n.markov@bsuir.by](mailto:a.n.markov@bsuir.by)

**Д.В. Орлов**

Студент 4 курса специальности “Вычислительные машины, системы и сети” БГУИР.

**С.Н. Нестеренков**

Кандидат технических наук, доцент, декан факультета компьютерных систем и сетей Белорусского государственного университета информатики и радиоэлектроники, доцент кафедры программного обеспечения информационных технологий. Автор публикаций на тему машинного обучения, алгоритмов принятия решений, искусственных нейронных сетей и автоматизации.

**А.Н. Марков**

Магистр технических наук, старший преподаватель кафедры ПИКС, заместитель начальника Центра информатизации и инновационных разработок Белорусского государственного университета информатики и радиоэлектроники.

**Аннотация.** Бинаризация – важный процесс в улучшении и анализе изображения. В настоящее время описано множество методов бинаризации. Эти алгоритмы создают бинарные аналоги цветных или представленных в оттенках серого изображений. Этот процесс крайне важен для нейронных сетей, которые направлены на распознавание образов в разных областях. В настоящей работе приводятся обширные подходы к бинаризации, различные методы, результат работы, плюсы и минусы.

**Ключевые слова:** бинаризация, глобальные методы, локальные методы, метод Оцу, метод Ниблэка, метод Саулова

**Введение.**

Двоичное представление изображения, черно-бело представление объекта и фона является предпочтительным для формата для анализа изображения документа, который состоит из текстовых объектов [1]. Следовательно, в этом случае бинаризация является важной предварительной обработкой изображения. Точность этого процесса влияет на выполнение последующих этапов обнаружения и извлечения текста из изображения, распознавание образов [2]. В конечном счете, этот процесс увеличивает видимость важной информации, отбрасывая ненужную.

Основной атрибут бинаризации – это порог  $T$ , со значением которого затем сравнивается яркость всех пикселей. После сравнения, пикселю присваивается одно из двух возможных значений: 0 – фон, 1 – передний план.

Общая схема бинаризации представлена на рисунке 1.

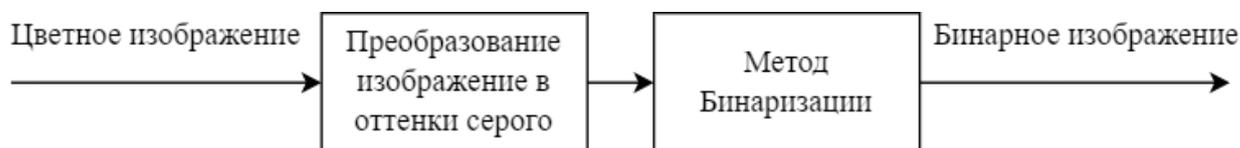


Рисунок 1 – Схема бинаризации

Существуют различные алгоритмы бинаризации, которые условно можно разделить на следующие группы, по признаку нахождения порога  $T$ : глобальные, локальные и гибридные.

1. Глобальные методы: в этих методах пороговое значение  $T$  рассчитывается для всего изображения и не изменяется в процессе.

2. Локальные методы: адаптивный метод нахождения порога  $T$  разделяет изображение на несколько областей. В каждой из этих областей рассчитывается свой порог. Значение порога формируется из набора пикселей, находящегося в этом регионе.

#### Глобальные методы.

Рассмотрим бинаризацию с нижним порогом, которая относится к глобальным методам. Порог  $T$  может рассчитываться по-разному: методом подбора к конкретной картинке, среднее значение между минимальным и максимальным уровнем яркости изображения. Рассмотрим на примере:



Рисунок 2 – Исходное изображение

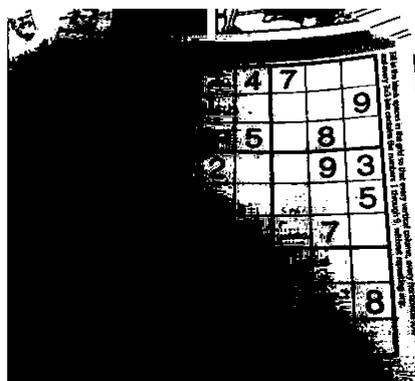


Рисунок 3 –Результат преобразования методом пороговой бинаризации

Одним из самых эффективных методов, относящихся к глобальной бинаризации, является метод Оцу [3].

Метод обрабатывает гистограмму изображения, сегментируя объекты путем минимизации дисперсии по каждому из классов. Обычно этот метод дает хорошие результаты для бимодальных изображений. Их гистограмма содержит два выраженных пика, представляющих разные диапазоны интенсивностей. Основной идеи является разделение гистограммы изображения на два кластера с порогом, определяемым в результате минимизации дисперсии этих классов. Полное уравнение для дисперсии может быть записано следующим образом:

$$\sigma_w^2(t) = w_1(t)\sigma_1^2(t) + w_2(t)\sigma_2^2(t) \quad [3]$$

где  $w_1(t)$ ,  $w_2(t)$  – вероятность двух классов, разделенных порогом  $T$ , значение которого лежит в диапазоне от 0 до 255 включительно.

В данном методе представлено два решения для нахождения порога. Первый это минимизировать внутриклассовую дисперсию  $\sigma_w^2(t)$ , второй это максимизировать межклассовую дисперсию используя выражение:

$$\sigma_b^2(t) = w_1(t)w_2(t) [\mu_1(t) - \mu_2(t)]^2$$

где  $\mu_i$  – среднее значение класса  $i$ .

Вероятность  $P$  рассчитывается для каждого значения пикселя в двух разделенных кластерах  $C_1$ ,  $C_2$  используя функции вероятностей

$$w_1(t) = \sum_{i=1}^t P(i)$$

$$w_2(t) = \sum_{i=t+1}^I P(i)$$

Следует отметить, что изображение может быть представлено в виде функции интенсивности  $f(x, y)$ , значения которой являются оттенками серого. Количество пикселей с заданным уровнем серого  $i$  обозначим через  $\square_i$ . Общее количество пикселей в изображении через  $n$ . Тогда вероятность появления оттенка серого  $i$ :

$$P(i) = \frac{\square_i}{n}$$

Значения интенсивностей для пикселей кластера  $C_1$  лежат в диапазоне  $[1, t]$ , а для  $C_2$  в диапазоне  $(t, I]$ , где  $I$  максимальное значение, равное 255 для серого изображения.

Следующий этап – это получение среднего значения кластеров  $C_1$ ,  $C_2$

$$\mu_1(t) = \sum_{i=1}^t \frac{\square_i \cdot i}{\square_1(t)}$$

$$\mu_2(t) = \sum_{i=t+1}^I \frac{\square_i \cdot i}{\square_2(t)}$$

Рассматривая уравнение, приведенное для внутриклассовой дисперсии, были найдены все компоненты для его подсчета

$$\sigma_1^2(t) = \sum_{i=1}^t [i - \mu_1(t)]^2 \frac{\square_i}{\square_1(t)}$$

$$\sigma_2^2(t) = \sum_{i=t+1}^I [i - \mu_2(t)]^2 \frac{\square_i}{\square_2(t)}$$

Следует помнить, что если порог выбран неправильно, то дисперсия одного из классов будет большой. Чтобы получить итоговую дисперсию нужно сложить внутриклассовую и межклассовую дисперсии.

$$\sigma^2_T = \sigma^2_w(t) + \sigma^2_b(t)$$

$$\sigma^2_b(t) = w_1(t)w_2(t)[\mu_1(t) - \mu_2(t)]^2$$

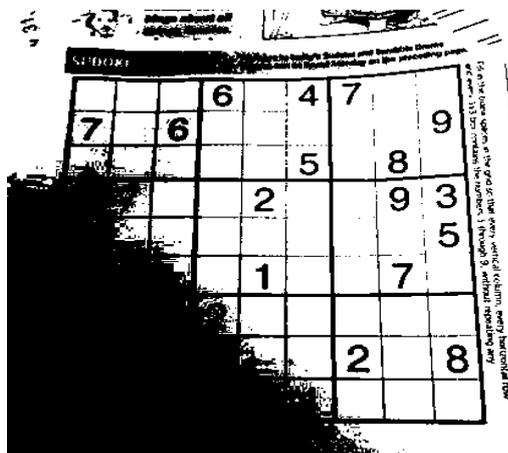


Рисунок 4 – Результат преобразования методом Оцу

Глобальные методы нахождения порога для бинаризации оказались, при простоте реализации и скорости алгоритмов, оказались весьма неплохи.

Однако на представленном примере часть информации была утеряна.

**Локальные методы.**

Метод Ниблэка: Алгоритм Ниблэка (1985 г.) вычисляет локальный порог для каждого пикселя с помощью скольжения прямоугольного окна по всему изображению. Порог T рассчитывается с использованием среднего  $m$  и стандартного отклонение  $\sigma$  всех пикселей в окне:

$$T = m + k * \sigma \quad [4]$$

где,  $m$  – среднее значение;  $\sigma$  – среднеквадратическое отклонение

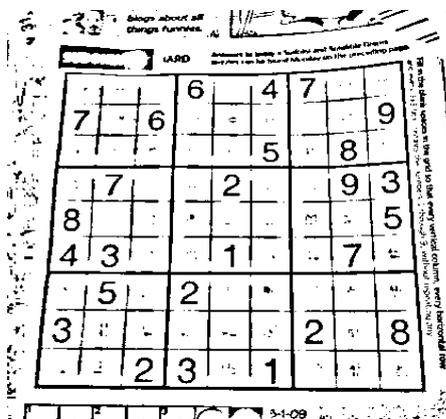


Рисунок 5 – Результат работы метода Ниблэка

Метод Саувола: Алгоритм Саувола это модификация метода Ниблэка (1997 г.), для повышения производительности бинаризации документов с фоном, содержащим легкую текстуру или разнообразное и неравномерное освещение.

$$T = m(1 - k(1 - \frac{\sigma}{\square})) [4]$$

где  $k$  лежит в диапазоне  $[0.2, 0.5]$  и диапазон интенсивности серого изображения  $R$ .

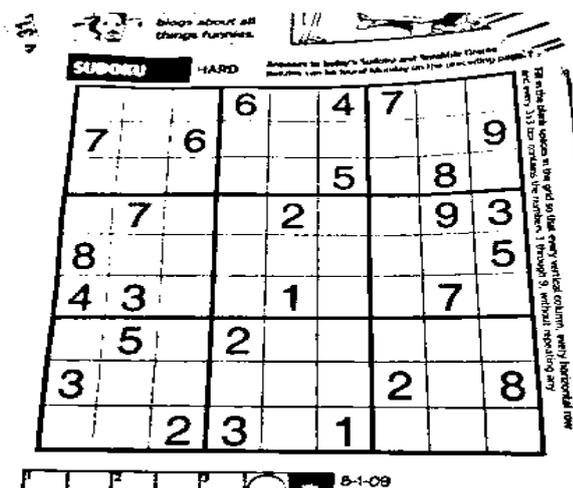


Рисунок 6 – Результат работы метода Саувола

Хотя глобальные методы проще и легче внедряются, они не способны справиться с некачественными изображениями (высокий уровень шум, низкая освещённость, ничтожная контрастность). С другой стороны, в методах локальной пороговой обработки пороговые значения определяются локально либо пиксель за пиксель. Основное преимущество локальных методов заключается в том, что в изображениях, где значительная часть фона зашумлена или переизбыток/недостаток освещения; существует множество пикселей, которые нельзя легко классифицировать как передний план или фон глобальными методами. В таких случаях локальная бинаризация более точна и уместна.

### Заключение

В этом исследовании мы рассмотрели несколько классов и принадлежащих к ним методов бинаризации изображений, в основном для преобразования уровней серого или цветного изображения в их двоичное представление. Представлено исходное изображение и результаты бинаризации для каждого из рассмотренных методов. Некоторые методы показали большую точность в сравнении с остальными [5]. К сожалению, до сих пор нету одного метода, который мог бы решить все проблемы изображения для получения его бинарного формата с большой точностью.

### Список литературы

- [1] Нестеренков, С.Н. Применение больших данных в электронном образовании / С.Н. Нестеренков, М.И. Макаров, Н.В. Ющенко, А.Д. Радкевич // BIG DATA and Advanced Analytics = BIG DATA и анализ высокого уровня : сб. материалов V Междунар. науч.-практ. конф. (Республика Беларусь, Минск, 13-14 марта 2019 года). В 2 ч. Ч. 2 / редкол. : В. А. Богуш [и др.]. - Минск : БГУИР, 2019. - С. 242-245.
- [2] Нестеренков, С. Н. Использование сверточных нейронных сетей для классификации и анализа тональности текстов / С. Н. Нестеренков, П. А. Федоров, В. А. Денисов // Информационные технологии и системы 2019 (ИТС 2019) : материалы междунар. науч. конф., Минск, 30 окт. 2019 г. / Белорус. гос. ун-т информатики и радиоэлектроники ; редкол.: Л. Ю. Шилин [и др.]. - Минск, 2019. - С. 248-249.
- [3] Rasche, C. Computer Vision / C. Rasche. – Bucharest : Polytechnic University of Bucharest, 2019. – 279 с.

[4] Peters, James F. Foundations of Computer Vision / James F. Peters. – Canada : Springer, 2017. – 443 с.

[5] Калоша, А.Л. Система анализа качества текстовых коллекций / А.Л. Калоша, М.А. Медунецкий, М.П. Хоронко, А.А. Александров, А.И. Гридасов, С.Н. Нестеренков // BIG DATA and Advanced Analytics = BIG DATA и анализ высокого уровня : сб. материалов VI Междунар. науч.-практ. конф. (Республика Беларусь, Минск, 20-21 мая 2020 года): в 3 ч. Ч. 2 / редкол. : В. А. Богуш [и др.]. - Минск : Бестпринт, 2020. - С. 369-375.

## **COMPARATIVE ANALYSIS OF IMAGE BINARIZATION METHODS**

***D.V. ORLOV***

*Student of BSUIR*

***S.N. NESTERENKOV***

*PhD, Associate professor,  
Dean of the Faculty of  
Computer Systems and  
Networks*

***A.N. MARKOV***

*Senior lecturer of the  
department, Deputy head of the  
Center for Informatization and  
Innovative Developments*

*Center for Informatization and Development of the Belarusian University of State Informatics and Radioelectronics,  
Republic of Belarus.*

*E-mail: [dimm4work@gmail.com](mailto:dimm4work@gmail.com), [s.nesterenkov@bsuir.by](mailto:s.nesterenkov@bsuir.by), [a.n.markov@bsuir.by](mailto:a.n.markov@bsuir.by)*

**Abstract.** Binarization is an important process in image enhancement and analysis. Currently, many binarization methods have been described. These algorithms create binary analogs of color or grayscale images. This process is extremely important for neural networks, which are aimed at pattern recognition in different areas. This paper presents extensive approaches to binarization, various methods, the result of the work, pluses and minuses.

**Keywords.** Binarization, global methods, local methods, Otsu method, Niblack method, Saulova method.