

УДК 004.912

ОСНОВНЫЕ МЕТОДЫ ОБРАБОТКИ ИЗОБРАЖЕНИЙ ПРИ ОФФЛАЙН-РАСПОЗНАВАНИИ РУКОПИСНОГО ТЕКСТА

Басанько А.С., Белов Ю.С.

*Калужский филиал ФГБОУ ВО «Московский государственный технический университет имени Н.Э. Баумана (национальный исследовательский университет)»,
Калуга, e-mail: antonbas1@rambler.ru*

В данной статье рассмотрены основные методы обработки изображения, которые применяются на этапе предварительной обработки при оффлайн-распознавании рукописного текста. Данные методы решают задачу повышения качества изображения, которая является достаточно актуальной, поскольку, чем лучше будет качество изображения, тем удобнее будет работать с ним на последующих этапах распознавания и, соответственно, качество распознавания также будет лучше. Рассмотрены основные этапы обработки изображения: преобразование изображения в градации серого, удаление дефектов изображения и отделение текста от фона. Представлены различные фильтры для удаления дефектов изображения: фильтр Гаусса (используется для подавления высокочастотного шума), медианный фильтр (используется для подавления шума «соль и перец») и фильтр на основе вейвлет-преобразования, который основан на представлении изображения в виде дискретного сигнала. Отмечены недостатки фильтров на основе вейвлет-преобразования. Рассмотрены основные алгоритмы отделения текста от фона: пороговая и адаптивная бинаризация. Отмечены основные недостатки этих методов. В рамках алгоритмов бинаризации рассмотрен метод Оцу, решающий задачу поиска порогового значения яркости. Также отмечены недостатки данного метода и способы их устранения.

Ключевые слова: распознавание рукописного текста, фильтр Гаусса, медианный фильтр, фильтр на основе вейвлет-преобразования, пороговая и адаптивная бинаризация, метод Оцу

BASIC IMAGE PROCESSING TECHNIQUES IN OFFLINE HANDWRITING RECOGNITION

Basanko A.S., Belov Yu.S.

*Kaluga branch of the Federal State Budget Education Institution of Higher Education
«Moscow State Technical University named after N.E. Bauman (National Research University)»,
Kaluga, e-mail: antonbas1@rambler.ru*

This article describes the basic methods of image processing that are applied at the stage of pre-processing in offline handwriting recognition. These methods solve the problem of increasing image quality, which is very essential because the better the image quality, the easier it will be to work with it in subsequent stages of recognition and, consequently, the quality of recognition will also be better. The basic stages of image processing: convert image to grayscale, remove imperfections from an image and separate text from background are described. A variety of filters for removing image defects: a Gaussian filter (used to suppress high frequency noise), median filter (used to suppress noise «salt and pepper»), filter based on wavelet transform, which based on the representation of the image in the form of a discrete signal are presented. The shortcomings of filter based on wavelet transform are noted. The main algorithms to separate text from the background: threshold and adaptive binarization are described. Their main shortcomings are noted. Method Otsu, which solve the task of finding threshold values of brightness, are described in the framework of the considered algorithms binarization. Also the shortcomings of this method and their solutions are noted.

Keywords: handwriting recognition, gaussian filter, median filter, filter based on wavelet transform, threshold and adaptive binarization, method Otsu

Оффлайн-распознавание рукописного текста обычно состоит из следующих этапов: предварительная обработка изображения, сегментация и нормализация, извлечение признаков, классификация и обработка результатов.

На этапе предварительной обработки изображения используются методы обработки изображений (фильтрация, шумоподавление и другие), и имеющие своей целью повысить качество изображения.

На данный момент задача распознавания рукописного текста является нерешенной. Задача повышения качества изображения при распознавании рукописного текста является

достаточно актуальной, поскольку чем лучше будет качество изображения, тем удобнее будет работать с ним на последующих этапах распознавания [1] и, соответственно, качество распознавания также будет лучше.

Улучшение качества изображения включает обычно в себя преобразование изображения в градации серого, удаление дефектов изображения и отделение текста от фона.

Преобразование изображения в градации серого

Обычно для упрощения дальнейшей работы с изображением делается преобразование изображения в градации серого [2].

Для каждого отдельного пикселя вычисляется значение яркости, которое измеряется в диапазоне от 0 до 255. Черный цвет соответствует 0 уровню яркости, а белый – 255 уровню. Это вычисление производится при помощи следующей формулы:

$$I = 0,299R + 0,587G + 0,114B,$$

где R, G, B – значения красного, зеленого и синего каналов соответственно.

Удаление дефектов

Удаление дефектов осуществляется стандартными методами обработки изображений. Наиболее часто для удаления шума используют фильтр Гаусса для подавления высокочастотного шума [3] и медианный фильтр для удаления шума «соль и перец» (рис. 1). Также перспективным фильтром является фильтр на основе вейвлет-преобразования [4].

Фильтр Гаусса

Данный фильтр основан на функциях Гаусса одной и двух переменных соответственно:

$$G(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{x^2}{2\sigma^2}},$$

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

где σ – «степень размытия» обработанного изображения (стандартное отклонение нор-

мального распределения); x, y – расстояние между пикселем (исходной точкой) и точкой, для которой подсчитывается значение функции Гаусса по вертикальной и горизонтальной оси соответственно [5].

Следовательно, при помощи функции Гаусса можно построить матрицу свертки, которая помогает для каждого пикселя изображения рассчитать средневзвешенное значение соседних пикселей:

$$r_{mn} = \sum_{i=-k}^k \sum_{j=-k}^k r_{m+i, n+j} G(i, j),$$

где k – размерность матрицы свертки.

Таким образом, посредством применения фильтра Гаусса шум будет подавлен, поскольку все зашумленные пиксели (яркость которых сильно отличается от яркости соседних пикселей) примут усредненное значение, в результате чего контуры объектов будут подчеркнуты, что положительно повлияет на результат распознавания образов (в том числе и рукописного текста) на цифровых изображениях.

Медианный фильтр

Понятие медианы лежит в основе медианного фильтра. Если множество A , состоящее из чисел A_i , где $i = 1..n$, отсортировано по возрастанию, то $A_{n/2}$ – является медианой этого множества [6]. Медиана будет делить отсортированный набор чисел на две части, где первая часть будет содержать числа, которые меньше, чем медиана, а вторая часть – больше.



а)



б)

Рис. 1. Примеры распространенных типов шумов: а) шум «соль и перец», б) высокочастотный шум

Распространенный способ реализации данного фильтра заключается в том, чтобы отсортировать значения яркостей пикселей при помощи окна [7] с нечетным радиусом и затем заменить значения яркостей пикселей на значение медианы результирующего множества.

Фильтр на основе вейвлет-преобразования

Так как изображение можно представить в виде дискретного сигнала, то для его обработки можно использовать фильтры, базирующиеся на частотном разделении в дискретной области. Вейвлет-анализ является достаточно перспективным способом анализа данных.

Сигнал можно представить следующим образом:

$$s(t) = f(t) + \sigma e(t),$$

где $f(t)$ – полезный сигнал, $e(t)$ – шум, σ – уровень шума, $s(t)$ – исследуемый сигнал.

Таким образом, вейвлет-преобразование позволяет удалить шум за 4 шага [8, 9]:

- разложение сигнала по базису вейвлетов;
- выбор порогового значения шума для каждого из уровней разложения;
- пороговая фильтрация коэффициентов детализации;
- восстановление сигнала.

Такой способ фильтрации лучше всего работает на гладких сигналах, т.е. на таких сигналах, в разложении которых лишь небольшое количество коэффициентов детализации значительно отличается от нуля.

Подбор вейвлета и глубины разложения обычно зависит от свойств фильтруемого сигнала.

Для выбора порога шума обычно используют критерии, которые минимизируют квадратичную функцию потерь для выбранной модели шума.

Данные фильтры реже используются по сравнению с медианными, потому что использование вейвлетов приводит к дополнительной параметризации программы и замедлению работы, поскольку требуется вычисление дополнительных массивов данных.

Отделение текста от фона

Отделение текста от фона является частным случаем задачи выделения объекта на изображении. Задача состоит в том, чтобы по изображению текста A построить бинарное изображение B , такое, что

$$B(i, j) = \begin{cases} 1, P(i, j) \in T_A \\ 0, P(i, j) \notin T_A, \end{cases}$$

где $P(i, j)$ – пиксель (i, j) , T_A – текст на изображении A .

Данное преобразование позволяет в дальнейшем использовать анализ связанных компонент, контуров, скелетов и т.д.

Наиболее часто используемым методом отделения текста от фона служит пороговая бинаризация (threshold binarization). Пусть дано изображение, $I(i, j)$ – яркость пикселя с координатами (i, j) . Пороговой бинаризацией изображения называется по-пиксельное преобразование $f(i, j)$, такое, что

$$f(i, j) = \begin{cases} 1, I(i, j) \geq d \\ 0, I(i, j) < d, \end{cases}$$

где d называется порогом бинаризации.

Обычно на гистограмме яркости изображения текста наблюдается два пика: высокий пик в области светлых пикселей, который соответствует фону, и более низкий пик в области темных, который соответствует тексту. Таким образом, задача поиска порогового значения яркости, т.е. такого, чтобы пиксели с яркостью выше этого значения (фон) будут считаться черными, а ниже (текст) – белыми (такое «инвертирование» цвета делается в целях упрощения применения многих алгоритмов в дальнейшем), является задачей поиска оптимального значения между двумя пиками гистограммы. Для решения этой задачи существуют хорошо изученный метод Оцу и его вариации.

Метод Оцу

В данном методе диапазон яркостей $[0; L]$ изображения делится на две части пороговым значением T . Суть алгоритма состоит в том, чтобы минимизировать внутриклассовую дисперсию, которая определяется как взвешенная сумма дисперсий двух классов. В алгоритме Оцу минимизация внутриклассовой дисперсии эквивалентна максимизации межклассовой дисперсии, которая рассчитывается следующим образом:

$$\sigma_b^2 = \omega_1 \omega_2 (\mu_1 - \mu_2)^2,$$

где σ_b – межклассовая дисперсия, w_1 и w_2 – вероятности первого и второго классов, μ_1 и μ_2 – средние арифметические значения каждого из классов. Каждая из перечисленных величин рассчитывается следующим образом:

$$\omega_1(t) = \sum_{i=1}^t p_i,$$

$$\omega_2(t) = \sum_{i=t+1}^L p_i = 1 - \omega_1(t),$$

$$\mu_1(t) = \sum_{i=1}^t \frac{ip_i}{\omega_0},$$

$$\mu_2(t) = \sum_{i=t+1}^L \frac{ip_i}{\omega_1}.$$

Схема алгоритма описывается следующим образом:

– Рассчитывается гистограмма на основе значений $p_i = n_i / N$, где N – суммарное количество пикселей изображения, n_i – количество пикселей с яркостью i .

– Со значения порога $t = 1$ делается проход по всей гистограмме, пересчитывая на каждом шаге дисперсию $\sigma_b(t)$. Если на каком-либо шаге дисперсия стала больше максимума, то обновляется максимум дисперсии и назначается новое текущее значение $T = t$.

– T – результирующее пороговое значение.

Недостатком данного метода является чувствительность к неравномерной освещенности. Для решения данной проблемы обычно получают компонент освещения путем низкочастотной фильтрации G изображения при помощи фильтра Гаусса [10].

Метод адаптивной бинаризации

Еще одним недостатком метода Оцу является слипание близко расположенных областей, что может повлиять на дальнейшую обработку и распознавание. Поэтому существует метод адаптивной бинаризации, который к тому же позволяет решить проблему разности освещенности.

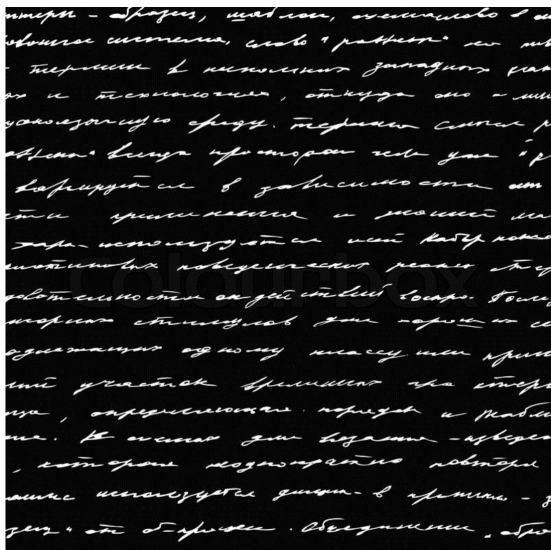


Рис. 2. Результат пороговой бинаризации

Для окрестности R пикселя вычисляется порог T . Порог T может являться средним

значением яркости по области R , медианой выборки из области R или вычисляться по формуле: $(I_{\max} - I_{\min}) / 2$. Значение пикселя $B(x, y)$ в бинарном изображении вычисляется следующим образом:

$$B(x, y) = \begin{cases} 1, & I(x, y) > T + C \\ 0, & I(x, y) \leq T + C \end{cases}$$

где $I(x, y)$ – яркость пикселя в исходном изображении, C – произвольная константа.

Результатом бинаризации изображения является бинарное изображение белого текста на чёрном фоне, соответствующего исходному изображению (рис. 2).

Таким образом, были рассмотрены основные методы обработки изображений при оффлайн распознавании рукописного текста. Следует грамотно подходить к выбору методов обработки изображения, поскольку при правильном выборе они являются залогом успешного распознавания текста.

Список литературы

1. Нестеров А.Ю., Бурмистров А.В., Белов Ю.С. Методы определения положения объекта в задачах распознавания образов // Электронный журнал: наука, техника и образование. – 2016. – № 2(6). – С. 82–89.
2. Гришанов К.М., Белов Ю.С. Методы выделения признаков для распознавания символов // Электронный журнал: наука, техника и образование. – 2016. – № 1(5). – С. 110–119.
3. Исследование и разработка метода распознавания рукописных цифр [Электронный ресурс]. – Режим доступа: <http://masters.donntu.org/2017/fknt/sokolov/diss/index.htm> (дата обращения: 25.03.2018).
4. Предварительная обработка изображений – Национальная библиотека им. Н.Э. Баумана [Электронный ресурс]. – Режим доступа: https://ru.bmstu.wiki/Предварительная_обработка_изображений (дата обращения: 25.03.2018).
5. Попов Г.А., Хрящёв Д.А. Об одном методе низкочастотной фильтрации гидролокационных изображений // Вестник АГТУ. – 2010. – № 1. – С. 63–68.
6. Фильтрация изображений [Электронный ресурс]. – Режим доступа: <https://studfiles.net/preview/5830083/page:4/> (дата обращения: 25.03.2018).
7. Логинов Б.М., Коржавый А.П., Белов Ю.С., Либеров Р.В. Методика распознавания и классификации образов структур многокомпонентных материалов для электронных систем // Электромагнитные волны и электронные системы. – 2017. – Т. 2, № 2(6). – С. 4–12.
8. Приложения вейвлет-анализа [Электронный ресурс]. – Режим доступа: <https://basegroup.ru/community/articles/wavelet-applications> (дата обращения: 25.03.2018).
9. Амелина В.Ю., Исаева В.С. Методы предварительной фильтрации изображения // Молодой учёный. – 2017. – № 20(154). – С. 179–182.
10. Попова В.В., Федосеев А.А. Предварительная обработка изображений в задаче оффлайн-распознавания рукописного текста // Научное сообщество студентов XXI столетия. Технические науки // Материалы IX студенческой международной заочной научно – практической конференции. – Новосибирск: СибАК, 2013. – С. 35–44.