

АДАПТИВНАЯ БИНАРИЗАЦИЯ ЦИФРОВЫХ ИЗОБРАЖЕНИЙ ТЕКСТОВЫХ БЛОКОВ

Введение

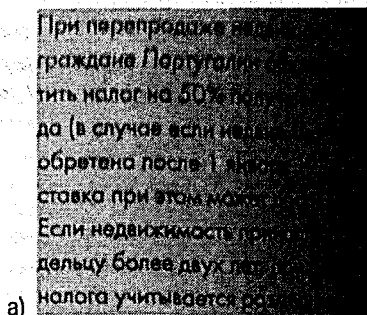
Бинарные изображения используются в самых разнообразных приложениях: обработка цифровых изображений документов, задачи промышленного зрения, медицинские исследования и др. При этом алгоритмы, для которых они являются входными данными, могут выполнять различные действия – от простых, например, извлечение характерных признаков областей, до гораздо более сложных, например, обнаружение и сопоставление объектов, распознавание образов и др.

Широта и удобство применения бинарных изображений усложняется весьма нетривиальной процедурой их получения. Существуют различные подходы, например, использующие глобальные пороги, локальные характеристики и др., однако, качество результата, ввиду значительной шумовой составляющей и яркостной неоднородности исходного изображения, может быть недостаточным для дальнейшего эффективного анализа. В частности, с такой проблемой приходится сталкиваться при бинаризации цифровых образов текстовых блоков, являющейся объектом исследования данной работы.

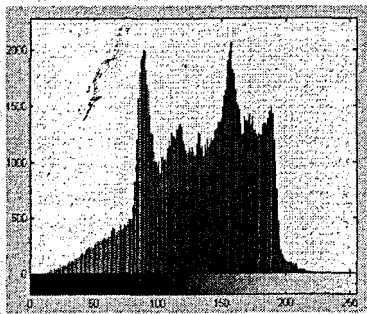
В качестве одного из вариантов решения указанной проблемы предлагается процедура адаптивной бинаризации, отличительной особенностью которой является использование контуров алфавитно-цифровых данных, образованных перепадами уровня яркости исходного изображения для проведения операции размывания реконструкцией. Целью применения данной операции является исключение фоновой составляющей из изображения текстовых блоков, что значительно облегчает итоговую пороговую бинаризацию.

Постановка задачи

В качестве входной информации будем рассматривать растровое полутоновое изображение, содержащее образы алфавитно-цифровых данных (символы, слова, предложения, абзацы) как, например, на рисунке 1(а). Результатом процесса обработки изображения является его бинарное представление, при этом элементы одного множества формируют сегменты текстовых образов, а остальные рассматриваются как фоновые.



а)



б)

полутоновое изображение (а), его гистограмма яркости (б)

Рисунок 1 – Пример текстового блока

Стандартные методы бинаризации не гарантируют получения качественных результатов, т.к. они исходят из предположения о бимодальности распределения значений яркости изображения, которое может существенно отклоняться от такой идеализированной модели, как, например, на рисунке 1(б).

- Предлагаемая процедура адаптивной бинаризации включает в себя следующие этапы:
- 1) выделение естественных характеристик областей интереса, в частности оценка высоты шрифта, получаемая в ходе анализа контуров изображения;
 - 2) исключение фоновой составляющей изображения, осуществляемое операцией размыкания реконструкцией, с использованием полученной на предыдущем этапе оценки;
 - 3) пороговая бинаризация обработанного изображения, учитывающая модальные значения гистограммы яркости.

Описание процедуры адаптивной бинаризации

Выделение естественных характеристик областей интереса

Первый из названных выше этапов связан с алгоритмами сегментации изображения текстового блока на различных уровнях и ввиду своей сложности является объектом отдельного исследования. Отметим лишь, что в их основе лежат полученные одним из детекторов (например, Кэнни) контуры полутонового изображения [1], подвергаемые кластеризации. Ее результатом являются обособленные последовательности текстовых образов (слова), с помощью которых и осуществляется оценка высоты шрифта.

Исключение фоновой составляющей изображения

Основное средство решения данной задачи – операция размыкания реконструкцией.

Морфологическая реконструкция представляет собой преобразование, в котором участвуют два изображения и один структурообразующий элемент. Одно из изображений, называемое маркером, является исходной точкой преобразования, другое изображение, называемое маской, накладывает определенные ограничения на отображение, используемый структурообразующий элемент определяет связность.

Если g – это маска, а f – маркер, то реконструкция g по f , которая обозначается $R_g(f)$, определяется следующей итеративной процедурой:

- 1) присвоить h_1 маркерное изображение f ;
- 2) построить структурообразующий элемент B ;
- 3) повторять: $h_{k+1} = (h_k \oplus B) \cap g$, до тех пор, пока не станет $h_{k+1} = h_k$.

При этом предполагается, что маркер f является подмножеством g : $f \subset g$.

Несмотря на простоту и удобство итеративной формулировки процесса реконструкции, существуют другие, более быстрые вычислительные алгоритмы, реализующие это преобразование. В частности, при проведении исследований в рамках данной работы была использована "гибридная реконструкция" [2].

Техника размыкания реконструкцией включает в себя следующие шаги:

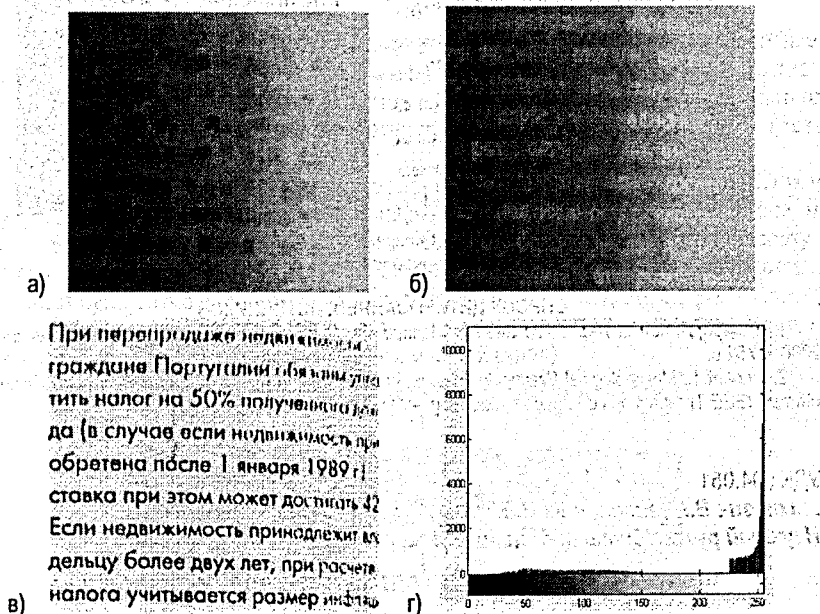
- 1) эрозия исходного полутонового изображения F : $A = F \ominus C$;
- 2) проведение морфологической реконструкции: $P = R_C(A)$;
- 3) исключение полученного изображения из исходного: $R = F \setminus P$.

Данные шаги по отношению к основной задаче второго этапа интерпретируются так:

- 1) удаление образов алфавитно-цифровых данных в изображении текстового блока;
- 2) восстановление фоновой составляющей изображения;
- 3) исключение фоновой составляющей из исходного изображения.

Основной проблемой в осуществлении указанных операций является подбор структурообразующего элемента эрозии S , позволяющего полностью удалить текстовые образы и при этом сохранить возможность восстановления фона алгоритмом реконструкции. В качестве источника оптимального размера элемента S предлагается использовать высоту шрифта h , оценка которого получена в ходе первого этапа.

В частности, в проводимых экспериментах использовался структурообразующий элемент, представляющий собой единичный квадрат размера $0.5^*h \times 0.5^*h$, который позволяет удалить текстовые образы и сохранить при этом локальные особенности яркости фоновой составляющей. Результаты выполнения операций второго этапа приведены на рисунке 2.



удаление текстовых образов (а), реконструкция (б) и вычитание фона (в), итоговая гистограмма (г)
Рисунок 2 – Исключение фоновой составляющей изображения текстового блока

Пороговая бинаризация изображения

Изображение, полученное на предыдущем этапе, отличается от исходного более выраженным контрастом текстовых образов и фона, при этом гистограмма распределения яркости значительно ближе к бимодальной модели. Следовательно, для получения итогового бинарного представления изображения можно воспользоваться одним из классических методов, например, алгоритмом Оцу [1].

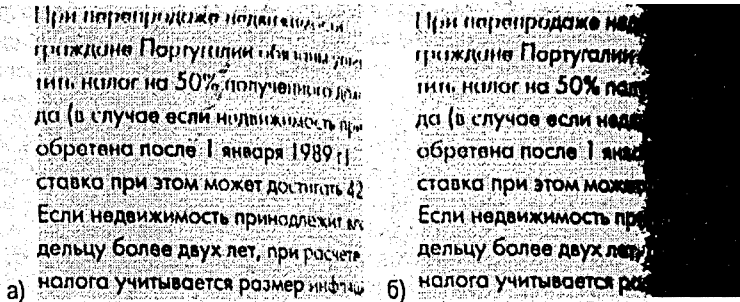
Данный алгоритм отличается эффективным вычислением глобального порога, минимизирующего внутригрупповую дисперсию двух групп пикселей, при этом значение порога располагается между двумя преобладающими модами (текста и фона) гистограммы распределения яркости изображения.

Итоговые результаты адаптивной бинаризации приведены на рисунке 3.

Выводы

В работе продемонстрирована эффективность процедуры адаптивной бинаризации цифровых изображений текстовых блоков, основанной на операции размыкания реконструкцией, являющейся мощным средством анализа изображений.

При этом обязательным условием получения качественного результата является учет естественных характеристик областей интереса, бинарное представление которых следует получить в ходе обработки.



Результат адаптивной бинаризации (а), алгоритма Оцу (б)
Рисунок 3 – Бинаризованное изображение текстового блока

СПИСОК ЦИТИРОВАННЫХ ИСТОЧНИКОВ

1. Шапиро, Л. Компьютерное зрение / Л. Шапиро, Дж. Стокман. – М.: БИНОМ. Лаборатория знаний. – 2006. – 752 с.
2. Vincent, L. Morphological Grayscale Reconstruction in Image Analysis: Applications and Efficient Algorithms // IEEE Transactions on Image Processing. – 1993. – Vol. 2, No. 2. – P. 176-201.

УДК 004.051

Михневич В.А., Никонович В.Б.

Научный руководитель: доцент Дунец А.П.

ПРИНЦИПЫ ОРГАНИЗАЦИИ АВТОМАТИЗИРОВАННОГО СКЛАДА

Введение

На заводах и фабриках хранится продукция. Магазины хранят свои товары, которые еще не попали на прилавок. Обычно такое место хранения называют складом. От увеличения помещения и количества груза соответственно усложняется обслуживание такого склада.

Представим себе оптовый магазин с большим складом. Допустим, это будет магазин компьютерной техники, в котором количество товаров может превышать 1000 или даже 10000 наименований устройств. Часто такие магазины не ограничиваются только компьютерными комплектующими, а продают также мобильные телефоны, ноутбуки, фотоаппараты, видеокамеры, расходные материалы и всякие другие подобные товары.

Другой пример – фармацевтические склады. Помимо большого количества разнообразных лекарств и препаратов, ситуация усложняется тем, что у таких товаров разные условия хранения (срок годности, температура, освещенность). Следовательно, приходится вести учет не только по наименованию товара, но и по другим параметрам. В