

СЕКЦИЯ 2. АЛГЕБРАИЧЕСКОЕ КОДИРОВАНИЕ И ОБРАБОТКА ИЗОБРАЖЕНИЙ

ОБНАРУЖЕНИЕ И СЕГМЕНТАЦИЯ ТЕКСТОВОЙ ИНФОРМАЦИИ РАСТРОВЫХ ИЗОБРАЖЕНИЙ

Н.Н. Кузьмицкий

В настоящий момент программное обеспечение для распознавания текстовой информации растровых изображений, основанное на технологиях Computer Vision, активно используется пользователями компьютеров. В то же время происходит постоянное расширение круга повседневных практических задач – как научных, так и производственных – для решения которых применяются данные технологии. При этом для достижения качественных результатов соответствующие специалисты должны уметь не только применять готовые программные продукты, но и понимать основные теоретические принципы их функционирования [1]. Кроме того, все более актуальной становится необходимость создания собственных законченных решений с учетом специфики конкретных задач (например, распознавания регистрационных номеров автомобилей). Подобный программный инструментарий неизбежно будет включать в себя ряд оригинальных алгоритмов и процедур, в первую очередь, для обнаружения текстовой информации на растровых изображениях, а также для сегментации и построения иерархической структуры текстовых образов.

В качестве входных данных рассматриваются растровые изображения, содержащие объекты следующих классов: документы, журнальные страницы, наружная реклама. Данные классы перечислены в порядке усложнения способов упорядоченности текстовой информации, разнородности стилистического оформления, фоновой и композиционной насыщенности. Основной целью обработки изображения является построение иерархической структуры текстовых образов, предполагающее их подробный анализ на каждом из следующих уровней: Колонка → Строка → Слово → Символ.

Поиск областей алфавитно-цифровых данных

Основное средство решения данной задачи – процедура контурной сегментации. Ее особенностью является одновременное использование двух базовых свойств яркости (разрывности и однородности), учет естественных свойств текстовых обра-

зов как основного объекта интереса, применение цветовой ориентации текстовых блоков: "темный текст + светлый фон" или "светлый текст + темный фон" [2].

Этапы обработки:

1) построение контурного представления полутонового изображения одним из известных средств, например, детектором Кэнни;

2) первый этап кластеризации контуров: его результатом являются обособленные текстовые образы ("слова");

3) второй этап кластеризации контуров: в результате его выполнения формируются крупные текстовых блоки ("колонки");

4) двухшаговая фильтрация кластеров: на основе аналитических (площадь, плотность и др.) и статистических (энтропия, гладкость и др.) признаков; из дальнейшего рассмотрения исключаются фоновые кластеры.

Далее текстовые блоки подвергаются процедуре адаптивной бинаризации, использующей оценку высоты шрифта текстовых образов. Данная оценка используется для проведения операции размыкания реконструкцией, целью которой является исключения фоновой составляющей [3].

Выделение алфавитно-цифровых последовательностей

Основным средством решения данной задачи являются алгоритмы сегментации текстового блока на строки и слова, использующие следующие признаки группировки алфавитно-цифровых данных: строки ориентированы горизонтально и обособлены некоторым межстрочным интервалом, а слова – некоторым межсловным интервалом.

Алгоритм сегментации текстового блока на строки:

1) текстовые сегменты бинарного изображения упорядочиваются по пространственному расположению: сверху вниз, слева направо;

2) для каждого текстового сегмента:

2.1) проводится поиск строки, с которой он имеет максимальное пересечение по высоте (в ходе анализа вертикальных координат минимальных прямоугольников);

2.2) если в результате поиска не было найдено подходящей строки, сегмент образует новую, иначе добавляется к найденной;

3) проводится дополнительное объединение строк:

3.1) если строки имеют пересечение по высоте более, чем половина высоты хотя бы одной из них, то происходит их слияние;

3.2) если высота строки или плотность ее сегментов, оказывается меньше определенных пороговых значений, то она сливается с одной из соседних строк.

Алгоритм сегментации строки на слова аналогичен предыдущему и основан на использовании пробельных регионов.

Извлечение алфавитно-цифровых образов

Основное средство решения данной задачи: алгоритм сегментации рукописного образа слова. Его основная особенность – использование секущих линий, определяемых исходя из аналитических характеристик кривизны формы символов [4].

Этапы обработки:

1) поиск опорных точек, определяющих возможные секущие отрезки: выполняется по признаку вогнутости внешнего контура изображения;

2) формирование множество возможных секущих: отрезки, задаются парами опорных точек либо одной опорной точкой и дополнительным условием, конкретизирующим направление;

3) в ходе последовательного перебора выбирается набор секущих, имеющих наибольшую интегральную оценку некоторого модуля распознавания.

Экспериментальная работа

На основе описанных алгоритмов с применением современных IT-технологий (.NET, Matlab и др.) был создан программный комплекс, позволяющий эффективно решать задачи обработки текстовой информации растровых изображений. В результате опробования комплекса на обширной выборке тестовых изображений надежность обнаружения текстовой информации составила 0,94...1 при степени подавления нетекстовых фрагментов 0,96...1. Проведенные исследования показали его сравнимость с известным аналогом ABBYY FineReader 9.0, при этом для изображений с наружной рекламой эффективность нового комплекса оказалась даже несколько большей.

В данной работе созданы новые процедуры обработки текстовой информации растровых изображений, объединенные в эффективные алгоритмы с последующей программной реализацией. Полученные результаты подтверждают перспективность создания специализированных средств решения научных и производственных задач в области Computer Vision.

Литература

1. *Шапиро Л.* Компьютерное зрение / Л. Шапиро, Дж. Стокман. – М.: БИНОМ. Лаборатория знаний. – 2006. – 752 с.
2. *Кузьмицкий Н.Н.* Контурная сегментация цифровых изображений документов / Н.Н. Кузьмицкий // Сб. конкурсных научных работ студентов и магистрантов. – Брест: БрГТУ, 2010. – Ч.1. – С. 278 – 282.
3. *Кузьмицкий Н.Н.* Адаптивная бинаризация цифровых изображений текстовых блоков / Н.Н. Кузьмицкий // Сб. конкурсных научных работ студентов и магистрантов. – Брест: БрГТУ, 2010. – Ч.1. – С. 283 – 286.
4. *Кузьмицкий, Н.Н.* Сегментация искаженного растрового изображения алфавитно-цифровой последовательности / Н.Н. Кузьмицкий // Сб. конкурсных научных работ студентов и магистрантов. – Брест: БрГТУ, 2009. – Ч.1. – С. 134 – 137.