

АЛГОРИТМ ВЫДЕЛЕНИЯ СТРУКТУРНЫХ СОСТАВЛЯЮЩИХ ДЛЯ РЕШЕНИЯ ЗАДАЧИ ОПТИЧЕСКОГО РАСПОЗНАНИЯ РУКОПИСНЫХ СИМВОЛОВ

Хаустов П.А., Максимова Е.И.

Томский политехнический университет, Институт кибернетики
exceibot@tpu.ru

Введение

Как известно, задача оптического распознавания символов является трудно формализуемой и имеет множество приближенных методов решений, каждый из которых обладает своими преимуществами и недостатками.

Большинство методов, которые применяются для решения задачи оптического распознавания символов, схожи с методами, которые используются для других задач классификации изображений. Как правило, такие методы предполагают использование искусственных нейронных сетей или аналогичных классификаторов, которые не позволяют отследить логику процесса классификации [1].

Если задача оптического распознавания символа не ограничивается только принятием решения о принадлежности образа к одному из классов, то для ее решения часто применяются методы на основе выделения структурных признаков. Структурные признаки обычно используются для выделения общей структуры образа. Они описывают геометрические и топологические свойства символа. Существует огромное количество способов задать некоторое графическое изображение с помощью набора структурных признаков, выделенных на этом изображении.

Как правило, если речь идет о выделении структурных признаков начертаний печатных или рукописных символов, то рационально использовать представление такого начертания в виде планарного графа. Способов представления начертания в виде такой модели, опять же, бесконечно много, как и алгоритмов получения каждой из такого рода моделей. Способы представления в виде планарного графа могут различаться степенью детализации и информативности.

Так, к примеру, ребро в различных моделях может описываться различными наборами величин. При тривиальном подходе для каждого ребра можно хранить лишь номера вершин, которое оно соединяет. В более сложных моделях ребро может описываться целым набором чисел, в котором, помимо соединяемых вершин и меры расстояния, находятся кривизна и даже целые множества графических примитивов для описания геометрической формы этого ребра.

Алгоритмы получения планарного графа для описания модели символа так же могут варьироваться от самых тривиальных,

опирающихся лишь на какие-то опорные точки графического начертания, алгоритмов до вероятностных алгоритмов попытки обнаружения каких-то особенностей начертания графических элементов.

К важным требованиям, которым должен соответствовать алгоритм решения задачи оптического распознавания символов на основе выделения структурных составляющих, относятся следующие требования:

1. Время работы алгоритма должно полиномиально зависеть от размеров исходного графического начертания символа.

2. Алгоритм должен быть детерминированным: для любого исходного графического начертания символа алгоритм должен за конечное время получать определенную модель представления этого символа.

3. Алгоритм должен обладать низким потреблением памяти. Количество необходимой для построения модели памяти не должно экспоненциально зависеть от размерности исходного начертания или итогового планарного графа.

4. Алгоритм должен выделять структурные признаки, опираясь только на входное изображение с графическим представлением начертания символа, не используя никаких дополнительных данных, в том числе и обучающей выборки.

В научной литературе можно найти множество попыток выделить структурные составляющие за полиномиальное время. Как правило, такие методы опираются на некоторую грамматику, с использованием которой задаются начертания различных символов. Однако, все большая степень повышения точности грамматики приводит к постоянным ее усложнениям.

Особый научный и даже практический интерес представляет группа методов, которые выделяют из растрового представления символа набор признаков достаточный, чтобы построить модель этого начертания на основе планарного графа, где ребра задаются целым набором графических примитивов.

Предложенный метод

Прежде всего, для получения информации о топологии начертания символа необходима предварительная скелетонизация его графического представления [2]. Так как каждый из общеизвестных алгоритмов скелетонизации, обладающих высоким быстродействием, имеет

свои недостатки, было решено последовательно использовать два известных алгоритма: алгоритм утончения Зонга-Суня и алгоритм Ву-Цая. Первый из этих алгоритмов периодически допускает присутствие на итоговом изображении неутонченных элементов, второй – зачастую удаляет небольшие элементы графического представления символа. Вследствие чего было высказано предположение об использовании алгоритма Ву-Цая для устранения нежелательных необработанных участков, оставшихся после использования алгоритма Зонга-Суня.

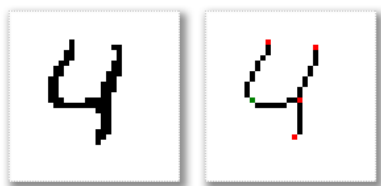


Рисунок 1 - Пример исходного изображения и результата его скелетонизации

Для получения топологической модели по уже утонченному изображению используется многократный запуск алгоритма Ли [3]. В роли вершин используются все пиксели, принадлежащие графическому представлению символа и не являющиеся фоном. С помощью набора эвристик обнаруживаются ключевые пиксели, которые последовательно удаляются перед очередным запуском алгоритма Ли, чтобы проанализировать расположение остальных пикселей относительно друг друга.

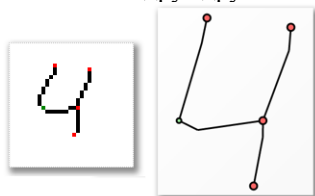


Рисунок 2 - Пример визуализации полученной модели

Ключевые пиксели подразделяются на две группы:

- пиксели, которые являются узловыми точками графического представления символа (такими точками являются места, где инструмент начертания поднимался, опускался или изменял направление своего хода);
- пиксели, которые являются точками перегиба, но не относятся к узловым точкам (такие точки используются для задания формы соединительных линий между ключевыми пикселями).

Все оставшиеся пиксели изображения относятся к соединительным элементам между ключевыми точками. Таким образом, графическое

представление символа можно преобразовать в планарный граф, имеющий композитные ребра, состоящие из обычных пикселей и пикселей – точек перегиба. Для уточнения структуры графа необходимо аппроксимировать каждый из участков черных пикселей между двумя точками-перегибами некоторым графическим примитивом. Тривиальным случаем такого графического примитива может служить отрезок или пара отрезков, образующих соединение. В некоторых случаях невозможно задать ребро с помощью одного-двух отрезков из-за высокой степени кривизны соединительной линии. В таком случае рационально использовать аппроксимацию с помощью дуги. При такой аппроксимации радиус и положение соответствующей итоговой дуге окружности должен определяться в зависимости от координат соединяемых ключевых точек и степени кривизны соединяющей линии. В некоторых случаях степень кривизны соединяющей линии соответствует слишком большому радиусу окружности, в результате чего окружность большей частью может выходить за пределы изображения, существенно отдаляя полученную топологическую модель от графического представления начертания символа. В таких случаях имеет смысл аппроксимировать такую соединяющую кривую с использованием дуги эллипса.



Рисунок 3 - Пример необходимости в использовании эллиптической дуги

Полученные результаты

При тестировании разработанного алгоритма установлено, что линейная зависимость быстродействия алгоритма от количества пикселей на изображении подтверждается высоким быстродействием на практике (построение топологической модели занимает не более секунды на современных процессорах).

Список использованных источников

1. Schantz, Herbert F., The history of OCR, optical character recognition – «Recognition Technologies Users Association», 1982. – 213
2. Роджерс Д., Алгоритмические основы машинной графики – М.: Мир, 1989. – с. 54-63
3. Кристофидес Н. Теория графов. Алгоритмический подход - М.: Мир, 1978. – 145 с.