

ГЕНЕТИЧЕСКИЙ АЛГОРИТМ ПОИСКА МНОЖЕСТВА КРИВЫХ ДЛЯ ОПТИЧЕСКОГО РАСПОЗНАВАНИЯ СИМВОЛОВ С ИСПОЛЬЗОВАНИЕМ МЕТОДА ПЕРЕСЕЧЕНИЙ

Хаустов П.А., Спицын В.Г., Максимова Е.И.

ФГБОУ ВПО «Национальный исследовательский Томский политехнический университет», Томск, Россия (634050, г. Томск, пр. Ленина, 30), e-mail: exceibot@tpu.ru

В данной статье описываются применение генетического алгоритма нахождения оптимального множества прямых для оптического распознавания символов на основе метода пересечений. В статье также описывается способ кодирования генетической информации об особи с использованием уравнения прямой с угловым коэффициентом. Для используемого варианта генетического алгоритма выбраны операторы кроссинговера и мутации, а также определены правила селекции. Для апробации алгоритма было реализовано приложение для оптического распознавания символов. Результатом работы является выявленная закономерность расположения набора прямых на графическом изображении символа для наилучшего качества распознавания. Для подтверждения сделанных выводов в статье производится визуализация полученных с использованием генетического алгоритма набора прямых для двух различных конфигураций генетического алгоритма.

Ключевые слова: генетические алгоритмы, оптическое распознавание символов, метод пересечений, алгоритм Брезенхэма, селекция, кроссинговер, мутация, функция приспособленности.

GENETIC ALGORITHM OF LINES SET SEARCHING FOR OPTICAL CHARACTER RECOGNITION BASED ON THE METHOD OF INTERSECTIONS

Khaustov P.A., Spitsyn V.G., Maksimova Y.I.

National Research Tomsk Polytechnic University, Tomsk, Russia (634050, Tomsk, Lenin Avenue, 30), e-mail: exceibot@tpu.ru

In this article the usage of genetic algorithm of optimal set of lines searching for optical character recognition based on the method of intersections has been proposed. The article also describes the method to encode the genetic information about an individual using the equation of a line with slope. For such genetic algorithm the rules of selection, operators of crossover and mutation have been proposed. For the algorithm testing the application for optical character recognition has been implemented. As a result, the consistent pattern for the set of lines have been found to maximize the optical character recognition rate. To confirm the made conclusion there is a visualization of a set of lines for different configurations of genetic algorithm.

Keywords: genetic algorithms, optical character recognition, method of intersections, Bresenham's algorithm, selection, crossover, mutation, fitness function.

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 [5].

Однако существуют методы, специально разработанные для бинаризованных изображений, и, в частности, изображений печатного текста. В таких методах для каждого пиксела известна его принадлежность к начертанию символа или фону. Особую роль среди

подобных подходов к решению задачи оптического распознавания символов играет группа методов, основанных на идее анализа пересечений графического представления символа с некоторым набором отрезков. Подобный подход получил название «метод пересечений» [1].

Идею метода пересечений можно интерпретировать огромным количеством способов – можно варьировать набор отрезков, способ геометрического представления символа, правила учета количества и общей длины пересечений каждого из отрезков с этим представлением.

Особый интерес вызывает выбор способа геометрического представления символа. Наиболее простым вариантом можно назвать анализ пересечений символа именно в исходном виде, но такой подход обладает некоторыми недостатками.

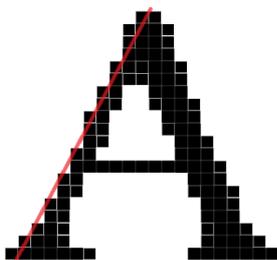


Рис.1. Пример нежелательного многократного пересечения отрезка и символа

Для случая, изображенного на рис. 1, характерно многократное пересечение отрезком графического представления символа. Фактически, группа пересекаемых пикселей до растеризации представляла собой отрезок, который не может иметь более одного интервала пересечения. Следовательно, такое представление символа зачастую искажает информацию о пересечениях символа с отрезком.

Наиболее точно описывает геометрию символа лишь его векторное представление. В общем случае символ будет представлен набором графических примитивов, которые будут ограничивать некоторые области изображения. Часть областей будет принадлежать символу, часть – фону.

При подобном представлении символа интервалы пересечения любого отрезка с начертанием этого символа можно определить максимально точно. Но в таком случае возникает проблема высокой вычислительной стоимости анализа интервалов пересечения. Для того чтобы избежать высокой вычислительной сложности и трудной программной реализации, необходимо некоторым образом перевести каждый из отрезков в растровый формат, что приведет к существенному упрощению анализа пересечений. Остается лишь выбрать способ перевода отрезков из векторного формата в растровый.

Описание алгоритма Брезенхэма. Существует множество способов перевести векторный отрезок в растровый формат изображения. Для того чтобы определить, какие пиксели двумерного растра необходимо отнести к изображению отрезка, требуется

анализировать их расположение относительно этого отрезка. Джек Е. Брезенхэм в 1962 году предложил способ определения принадлежности каждого из пикселей двумерного растра отрезку с использованием лишь целочисленной арифметики [3].

В ходе алгоритма предполагается проход по каждому значению координат одной из осей и определение значения координаты на другой оси для размещения очередного пиксела.

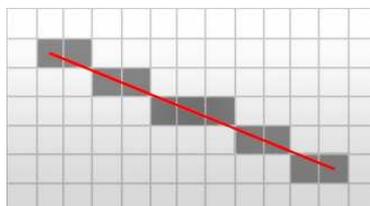


Рис. 2. Иллюстрация работы алгоритма Брезенхэма

Для того чтобы определить наиболее подходящее значение на другой оси координат, необходимо выбрать наиболее близкий к отрезку пиксел двумерного растра. Такую задачу можно решить, используя только целочисленную арифметику, с помощью уравнения прямой в отрезках.

Алгоритм Брезенхэма в методе пересечений. При использовании растрового представления отрезков в методе пересечений анализ интервалов наложения изображения отрезка на изображение символа существенно упрощается. Наиболее простым из допустимых вариантов оценки меры пересечения является тривиальный подсчет количества общих пикселей у двух растровых представлений. В таком случае геометрические потери при растеризации гораздо меньше влияют на результат распознавания. Для более точного описания геометрических потерь можно воспользоваться иллюстрацией.

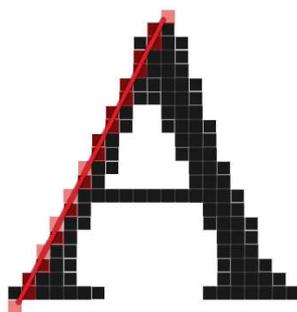


Рис. 3. Пример геометрических потерь

На рис. 3 оттенками красного цвета показаны пиксели, принадлежащие растеризации изображенного отрезка. Темный оттенок соответствует пикселям, которые относятся к графическому представлению символа, светлый оттенок соответствует остальным пикселям растрового представления отрезка. Как можно заметить, темно-красные пиксели в виду некоторой погрешности растрового представления образуют множество непрерывных интервалов наложения, что при точном геометрическом представлении соответствовало бы лишь одному интервалу наложения. При статистическом подсчете количества общих точек

такая погрешность не выглядит столь существенной, сколь может оказаться при более детальном геометрическом анализе. Для приведенного на рис. 3 примера вместо одного интервала наложения можно увидеть сразу пять интервалов, что может привести к затруднениям при проведении анализа количества непрерывных пересечений.

Теперь можно описать алгоритм оптического распознавания символов на основе метода пересечений и алгоритма Брезенхэма.

Шаг 1. Сформировать некоторый набор отрезков.

Шаг 2. Получить растровое представление каждого из отрезков, используя алгоритм Брезенхэма.

Шаг 3. Сформировать набор изображений символов – шаблонов, с которыми будет осуществляться сравнение.

Шаг 4. Для каждого изображения-шаблона произвести наложение каждого из отрезков набора. Для каждого из таких наложений O_i запомнить количество общих пикселей C_i . В дальнейшем вектор значений $V = \{C_1, C_2, \dots, C_k\}$ будет использоваться для представления данного изображения – шаблона.

Шаг 5. Для каждого из изображений, класс которого необходимо определить, описанным в шаге 4 образом получить вектор $V = \{C_1, C_2, \dots, C_k\}$. Найти наиболее похожий вектор в наборе шаблонов, определить класс, к которому относится символ, соответствующий этому вектору. Отнести к этому классу данное изображение.

Поиск набора прямых с помощью генетического алгоритма. Задача выбора оптимального набора прямых для решения задачи оптического распознавания символов с использованием метода пересечений является трудно формализуемой и существенно зависит от типа графического представления символов, размеров изображения, а также от набора шаблонных изображений символов.

Одним из способов нахождения оптимального набора прямых является использование генетического алгоритма [2, 4]. Для применения генетического алгоритма необходимо выбрать способ кодирования генетической информации, методы кроссинговера и мутации, способ селекции и функцию приспособленности.

Наиболее простым способом описания прямой являются уравнения прямой с угловым коэффициентом:

$$y = kx + b, k = \operatorname{tg}(\varphi)$$

Существенным недостатком такого представления прямой является невозможность описания вертикальной прямой таким образом. Для описанного метода пересечений такой недостаток не критичен, ведь в роли вертикальной прямой может выступать прямая, сколь угодно близкая к вертикальной. При растеризации методом Брезенхэма такая прямая будет

неотличима от вертикальной, и, следовательно, можно считать, что уравнение прямой с угловым коэффициентом позволяет задавать прямую для предложенного метода пересечений без потерь. Основным преимуществом такого описания прямой является необходимость хранить лишь два вещественных числа – значения углового коэффициента прямой и свободного члена. Следовательно, хранить набор из n прямых для метода пересечений можно с использованием лишь $2n$ вещественных чисел.

Как уже ранее было отмечено, хранить набор прямых можно с использованием $2n$ вещественных чисел. Следовательно, для хранения отдельной особи популяции рационально использовать вещественное кодирование. Для описанного способа вещественного кодирования можно использовать двухточечный или одноточечный операторы кроссинговера. Существенным недостатком двухточечного оператора кроссинговера можно назвать его сравнительно высокую разрушительную способность. Также двухточечный кроссинговер требует чуть более существенных вычислительных затрат. В роли оператора мутации в таком случае может выступать изменение одной из хромосом на некоторую малую величину. Другими словами, в процессе мутации с заданной вероятностью на некоторую небольшую величину может измениться либо угловой коэффициент, либо свободный член одной из прямых набора.

В качестве метода селекции был выбран рулеточный метод селекции, при котором вероятность участия особи в кроссинговере пропорциональна значению его функции приспособленности. В качестве функции приспособленности можно использовать долю успешно распознанных символов некоторого оценочного набора.

Эксперименты и полученные результаты. Изначально для получения оптимального набора прямых был использован генетический алгоритм со следующими значениями параметров:

- количество особей в популяции равно 16;
- количество прямых у одной особи равно 8 (16 хромосом);
- количество эпох эволюции равно 20.

На стадии инициализации начальной популяции был получен набор случайных особей, затем они были отсортированы в порядке убывания значений функции приспособленности. Набор прямых, соответствующий лучшей особи, можно назвать исходным набором прямых. После того, как было осуществлено 20 эпох эволюции, был получен окончательный набор особей, который также был отсортирован по убыванию значений функции приспособленности. Набор прямых, соответствующий лучшей особи, в таком случае можно назвать окончательным набором прямых. На изображении ниже приведены исходные и окончательные наборы прямых для трех различных запусков.

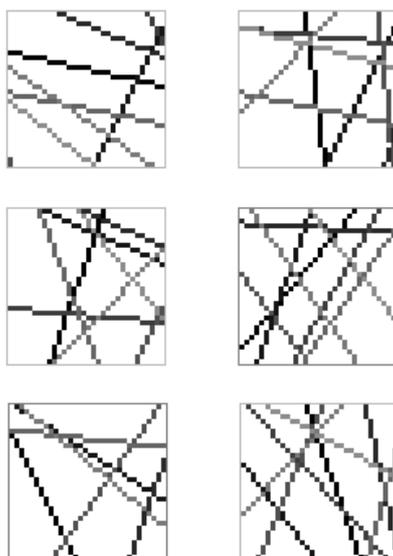


Рис. 4. Исходные (слева) и соответствующие им окончательные (справа) наборы прямых

Аналогичным образом был апробирован генетический алгоритм со следующими значениями параметров:

- количество особей в популяции равно 32;
- количество прямых у одной особи равно 16 (32 хромосомы);
- количество эпох эволюции равно 20.

Исходные и окончательные наборы прямых представлены на рис. 5:



Рис. 5. Исходные (слева) и соответствующие им окончательные (справа) наборы прямых

Выводы. В результате проделанной работы был выявлен ряд закономерностей формирования оптимального набора прямых. По окончании работы генетического алгоритма образуется набор прямых, который заполняет наибольшее количество пространства в центральной области изображения. То есть набор прямых изменяется так, чтобы среднее количество пересечений с начертанием символа было максимально. Важно заметить, что в результате эволюции особи с регулярной структурой не выживают, уступая место особям, у которых интервал различных углов наклона прямых достаточно обширен.

Список литературы

1. Роджерс Д. Алгоритмические основы машинной графики. – М.: Мир, 1989. – [С. 54-63].
2. Цой Ю.Р. О математических моделях эволюционных алгоритмов // Перспективные информационные технологии и системы. – 2006. – № 2. – [С. 42-47].
3. Bresenham, J. E., Algorithm for computer control of a digital plotter – «IBM Systems Journal 4», 1965. – [P. 25-30].
4. Janikow C. Z., Michalewicz Z., An Experimental Comparison of Binary and Floating Point Representations in Genetic Algorithms – «Proceedings of the Fourth International Conference on Genetic Algorithms», 1991. – [P. 31-36].
5. Schantz, Herbert F., The history of OCR, optical character recognition – «Recognition Technologies Users Association», 1982. – 213 p.

Рецензенты:

Ким В.Л., д.т.н., профессор кафедры вычислительной техники Института Кибернетики ФГБОУ ВПО НИ ТПУ, г. Томск;

Авдеева Д.К., д.т.н., профессор, директор ООО «Медприбор», г. Томск.