

3. Международный форум «Клиническая иммунология и аллергология – междисциплинарные проблемы» Бронхиальная астма: современные возможности контроля. Медицинский портал для врачей Umedp. URL: [http://umedp.ru/articles/mezhdunarodnyy\\_forum\\_klinicheskaya\\_immunologiya\\_i\\_allergologiya\\_mezhdistsiplinarnye\\_problemy\\_bronkhi.html](http://umedp.ru/articles/mezhdunarodnyy_forum_klinicheskaya_immunologiya_i_allergologiya_mezhdistsiplinarnye_problemy_bronkhi.html) (Дата обращения: 10.02.2016)
4. Берестнева О.Г., Шаропин К.А., Старикова А.В., Кабанова Л.И. Технология формирования баз знаний в медицинских и информационных системах // Известия Южного федерального университета. – 2010. – Т. 109. - №8.

## АЛГОРИТМ ОПТИЧЕСКОГО РАСПОЗНАВАНИЯ РУКОПИСНЫХ СИМВОЛОВ НА ОСНОВЕ ПОСТРОЕНИЯ СТРУКТУРНОЙ МОДЕЛИ

*Хаустов П.А., Спицын В.Г.*  
(г. Томск, Томский политехнический университет)  
e-mail: [exceibot@tpu.ru](mailto:exceibot@tpu.ru)

## ALGORITHM FOR OPTICAL HANDWRITTEN CHARACTERS RECOGNITION BASED ON STRUCTURAL MODEL COMPOSING

*Khaustov P.A., Spitsyn V.G.*  
(Tomsk, Tomsk Polytechnic University)

The algorithm for character topology composition has been proposed. Metrics for character graphs comparison has been suggested. The proposed algorithm has been implemented in character processing application and has been approved on MNIST handwriting characters database and writing characters examples from the forms of a unified state exam.

**Keywords:** structural model, structural components, optical character recognition, ocr.

**Введение.** Задача оптического распознавания символов актуальна при разработке большого количества интеллектуальных систем [1]. Важный класс алгоритмов, имеющих особенно высокую точность, основывается на использовании структурных моделей. Разработка метода такого класса представляет как научный, так и практический интерес.

**Предложенный метод.** Одним из вариантов представления начертания символа является его представление в виде планарного графа, вершинами которого являются некоторые ключевые точки начертания символа, а ребрами – соединяющие их участки пикселей этого начертания. Каждое из таких ребер может быть задано в виде некоторого количества последовательно соединенных отрезков, дуг и, возможно, эллиптических дуг.

Для получения информации о топологии начертания символа необходима предварительная скелетизация (утонение) его графического представления. Так как каждый из общеизвестных алгоритмов скелетизации, обладающих высоким быстродействием [3], имеет свои недостатки, было решено последовательно использовать два известных алгоритма: алгоритм утончения Зонга-Суня [2] и алгоритм Ву-Цая [3].

Для получения топологической модели по уже утонченному изображению используется многократный запуск алгоритма Ли, который является частным случаем применения алгоритма обхода в ширину для компонент восьми-связности. В роли вершин используются все пиксели, принадлежащие графическому представлению символа и, при этом, не являющиеся фоном.

Далее необходимо выделить на скелетизированном изображении ключевые точки. Такие точки будут являться вершинами графа итоговой модели, а соединяющие их ребра будут

формироваться из оставшихся пикселей. На первый взгляд к ключевым точкам можно отнести лишь пиксели, которые можно определить, основываясь на количестве и расположении соседних с ними пикселей. Однако есть точки, которые нельзя отнести к ключевым, но проходящее через них соединяющее ребро существенно изменяет свое направление в этих пикселях. Такие точки будем называть изгибами.

В ходе алгоритма поиска изгибов, будем считать изгибы ключевыми точками. Перед запуском алгоритма необходимо выделить все ключевые точки. После работы алгоритма все ключевые точки, которые были найдены после этой стадии, будут считаться изгибами. На каждой итерации алгоритма будем запускать алгоритм Ли. Запуск будем производить одновременно из всех найденных к этому моменту ключевых точек. Две встречных волны обхода будут соединяться в точке внутри некоторого соединяющего ребра. Если на пути к этой точке хотя бы одна из волн изменяла свое направление более чем на 60 градусов, то будем считать точку, в которой это произошло, ключевой. Если на очередной итерации не найдено ни одной новой ключевой точки, алгоритм заканчивает свою работу, все ключевые точки, найденные в ходе выполнения описанных итераций, полагаются изгибами.

Все пиксели изображения, кроме ключевых и изгибов, относятся к соединительным элементам между ключевыми точками. Таким образом, графическое представление символа можно преобразовать в планарный граф. Для упрощения структуры графа необходимо аппроксимировать каждый из участков черных пикселей между двумя точками-перегибами некоторым графическим примитивом: отрезком, дугой или эллиптической дугой.

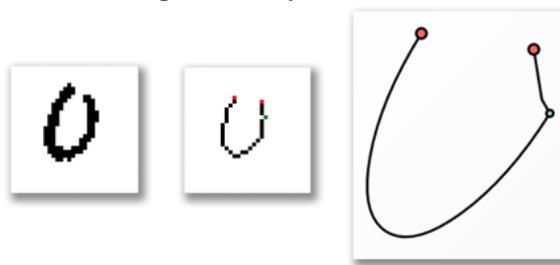


Рисунок 1 – Пример необходимости использования дуги эллипса в топологическом представлении

Для проверки некоторого начертания символа на соответствие указанному шаблону необходимо ввести некоторую функцию, отражающую степень сходства между двумя топологическими моделями. Например, оценкой степени схожести двух соединяющих линий может послужить площадь фигуры, заключенной между ними: чем меньше эта площадь – тем больше похожи эти линии.

После чего можно составить двудольный граф. Вершины первой доли этого графа будут соответствовать ребрам первой структурной модели, а вершины второй доли – ребрам второй модели. Каждое из ребер в полученном графе будут соответствовать определенной паре соединительных линий двух сравниваемых топологических моделей. Весом ребра в таком случае будет площадь фигуры, заключенной между двумя соответствующими этому ребру кривыми. В таком графе требуется найти максимальное паросочетание минимального веса. Такая задача решается с помощью алгоритма, похожего на алгоритм Эдмонса-Карпа поиска максимального потока.

**Полученные результаты.** Предложенный алгоритм был реализован в виде консольного приложения на языке программирования C++. Исследование быстродействия и качества распознавания алгоритма выполнялось на общеизвестном наборе рукописных цифр MNIST.

Таблица 2 – Результаты сравнения качества распознавания на наборе MNIST

	SM	IM	SVM	PNN
--	----	----	-----	-----

E = 3	93.2	87.4	83.8	70.9
E = 5	95.1	88.7	85.5	73.2

Как можно заметить, в условиях малого количества эталонных изображений, предложенный алгоритм существенно превосходит свои аналоги. Сравнительно низкие результаты нечетких классификаторов можно объяснить тем, что для качественного их обучения требуется достаточно большой размер обучающей выборки.

#### ЛИТЕРАТУРА

1. Schantz, Herbert F., The history of OCR, optical character recognition – «Recognition Technologies Users Association», 1982. – 213 p.
2. T.Y. Zhang, C.Y. Suen, A Fast Parallel Algorithm for Thinning Digital Patterns, Image Processing and Computer Vision. 3, 1984, pp. 236-239.
3. R.-Y. Wu, W.-H. Tsai, A new one-pass parallel thinning algorithm for binary images, Pattern Recognition Letters. 13, 1992, pp. 715-723.