

УДК 004.89

Н.М. Кораблев, А. В. Легедина, О.Г. Лебедев

Харьковский национальный университет радиоэлектроники, Харьков

СКЕЛЕТОНИЗАЦИЯ СИМВОЛОВ С ИСПОЛЬЗОВАНИЕМ МОДЕЛИ КЛОНАЛЬНОЙ СЕЛЕКЦИИ

В статье рассматривается задача подготовки изображения для распознавания, а именно, задача скелетонизации символов. Проведен анализ существующих методов решения задачи скелетонизации и предложен гибридный подход, который основан на использовании искусственных иммунных систем, а именно, модели клональной селекции, и алгоритма скелетонизации Зонге-Суня. Проведены экспериментальные исследования и определены дальнейшие направления исследований, ориентированные на решение задачи распознавания символов.

Ключевые слова: искусственная иммунная система, распознавание символов, клональная селекция, аффинность, антитело, антиген, скелетонизация, алгоритм Зонге-Суня.

Введение

В настоящее время разработано и реализовано большое количество методов распознавания образов. Распознавание образов можно разделить на три основных этапа [1]: 1) подготовка изображения (предварительная обработка); 2) выделение признаков на изображении; 3) отнесение символа на основе выделенных признаков к тому или иному образу (сама задача распознавания). В статье рассматривается этап подготовки изображения, а именно построение скелета (остова) распознаваемого символа.

Построение остова является важным этапом в задачах распознавания символов. В связи с этим большое значение имеет корректность его представления – соответствие структуры рассматриваемого объекта и его устойчивость к искажениям формы (зашумленность, разрывы, нечеткость изображения др.). Существующие подходы для решения указанной задачи скелетонизации (выделение каркаса монохромного изображения с шириной линий в один пиксель) имеют свои недостатки, основными из которых являются наличие разрывов при построении остова, наличие избыточных точек, утолщающих кривую остова и др. [2, 3].

Среди парадигм искусственного интеллекта, таких как искусственные нейронные сети, нечеткая логика, генетические алгоритмы и др., позволяющих автоматизировать процессы распознавания символов, в последнее время активно используются искусственные иммунные системы (ИИС) [4, 5], которые характеризуются высоким уровнем адаптивности, обладают способностью к обобщению и самообучению, имеют распределенный характер вычислений др.

Задача распознавания символов с использованием ИИС является перспективной, поскольку одна из основных задач биологической иммунной систе-

мы – разделение объектов на классы, а именно на «свои» (self) и «чужие» (non-self) клетки. При этом существует возможность обучения с помощью экзепляров только одного класса. Используя иммунную память, иммунная система поддерживает идеальный баланс между экономией ресурсов и выполнением своих задач за счет сохранения минимально необходимых знаний, получаемых в процессе обучения. Поскольку формирование иммунного ответа всегда происходит случайным образом, существует возможность определять инородные тела даже в тех случаях, когда не существует явных представлений об их структуре [4, 5].

Целью работы является разработка гибридного метода построения скелета на основе алгоритма Зонге-Суня [2] и модели клональной селекции [4, 5].

Постановка задачи

Понятие скелета символа широко используется в теории распознавания образов, так как позволяет получить наиболее полную информацию о геометрии объекта, хорошо отображая его форму [1]. Скелет символа P , согласно [1, 2], можно рассматривать как множество σ его внутренних точек, таких, что каждая точка $p \in \sigma$ равноудалена, по крайней мере, от двух точек на границе символа P .

Формулируемая задача скелетонизации [6] состоит в нахождении такого изображения, размеры которого будут равны размерам исходного изображения (высота и ширина скелета должны быть равны высоте и ширине исходного изображения), т.е.:

$$H_{oi} = H_{ii}, \quad W_{oi} = W_{ii}, \quad (1)$$

$$T_{oi} = \begin{cases} T_{ii}, & \text{если } T_{ii} = 1 \\ 1, & \text{если } T_{ii} \neq 1 \end{cases}, \quad (2)$$

где H_{oi}, H_{ii} – высота выходного и исходного изображений (скелета) соответственно, W_{oi}, W_{ii} – ши-

рина выходного и исходного изображения соответственно, T_{oi}, T_{ii} – толщина выходного и исходного изображения соответственно. Требуемый скелет будет равен исходному изображению в том случае, если толщина исходного изображения равна одному пикселю; если толщина исходного изображения более чем 1 пиксель, то выполняется скелетонизация изображения согласно предложенному ниже алгоритму. Длина и ширина исходного изображения получаются на основе максимально отдаленной пары крайних точек скелета символа P .

Метод решения

Одним из наиболее эффективных алгоритмов построения скелетов бинарных изображений является алгоритм Зонге-Суня [2, 6], который используется для формирования первоначального варианта скелета. Алгоритм Зонге-Суня состоит в следующем: обозначим точки изображения, принадлежащие символу, через 1, а точки, принадлежащие фону, через 0. Тогда для каждой точки p контура объекта вводится матрица, описывающая её окрестность:

$$P = \begin{pmatrix} p_1 & p_2 & p_3 \\ p_8 & p_0 & p_4 \\ p_7 & p_6 & p_5 \end{pmatrix}. \quad (3)$$

Точка p считается кандидатом на удаление при выполнении следующих условий:

$$2 \leq \sum_{i=1}^8 p_i \leq 6; \quad (4)$$

$$A(p_i) = 1, \quad (5)$$

$$p_2 p_4 p_6 = 0; \quad (6)$$

$$p_4 p_6 p_8 = 0, \quad (7)$$

где $A(p_i)$ – число переходов из 0 в 1 последовательности $p_1, p_2, p_3, p_4, p_5, p_6, p_7, p_8$, замыкая цепочку на p_1 , т.е. вокруг пикселя p_0 существует только один переход от 0 к 1. Выполнение условия (4) гарантирует, что рассматриваемая точка не является конечной. Условие (5) нарушается в том случае, если p находится на участке контура толщиной в один пиксель и её удаление в этом случае привело бы к нарушению скелета. Условия (6) и (7) позволяют удалить юго-восточные точки, так как их одновременное выполнение возможно при $p_2 = 0, p_8 = 0, p_6 = 0$.

Таким образом, на первом шаге помечаем для удаления точки контура, для которых выполняются условия (4) – (7), при этом удаление как таковое не производится, чтобы не нарушалась структура объекта. Удаление помеченных точек производится только после проверки всех точек границы. Аналогично выполняется второй шаг, удаление северо-западных точек, но условия (6), (7) заменяются на следующие:

$$p_2 p_4 p_8 = 0, \quad (8)$$

$$p_2 p_6 p_8 = 0. \quad (9)$$

Процедура повторяется, пока не прекратится процесс удаления точек.

На практике данный метод не всегда дает достаточно точный результат, и полученный скелет требует дополнительной обработки для дальнейшего выделения признаков: удаление коротких линий, объединение близких точек, исключение малых внутренних контуров. Для получения более точного скелета изображения, по которому можно сразу же выделить признаки, в работе предлагается гибридный метод, объединяющий алгоритм Зонге-Суня с иммунным подходом.

Среди основных моделей ИИС выделяют модель иммунной сети, модель клональной селекции и модель отрицательного отбора, каждая из которых имеет свои преимущества и недостатки [4, 5]. Для предлагаемого гибридного метода построения скелета была выбрана модель клональной селекции, которая обладает наилучшими качествами для решения поставленной задачи. А именно, возможность исследования неопределенной и неточной информации, возможность работы с нелинейностью, возможность корректировки результатов [4, 5].

Предложенный в работе метод, объединяющий алгоритм Зонге-Суня с моделью клональной селекции, позволяет получить более точный скелет изображения, по которому можно сразу выделить признаки. Гибридный подход заключается в следующем. ИИС в данном случае представляет собой популяцию антител, где каждое антитело соответствует пикселю (клетке) на изображении и представляет собой строку фиксированной длины [4, 5]:

$$Ab = \left\langle \begin{matrix} x_1 \dots x_i \dots x_m, y_1 \dots y_i \dots y_m, \\ c_1 \dots c_i \dots c_m, s_1 \dots s_i \dots s_m \end{matrix} \right\rangle, \quad (10)$$

где $x_1 \dots x_i \dots x_m, y_1 \dots y_i \dots y_m$, $i = \overline{1, n}$ – координаты i -й точки из популяции размером n антител, каждая из которых кодируется m разрядами, $c_1 \dots c_i \dots c_m$ – цвет данной точки, $s_1 \dots s_i \dots s_m$ – выживаемость клетки, устанавливаемая алгоритмом Зонге-Суня – {0;1} (избыточна клетка или нет). Антигены Ag представляют собой чужеродные объекты, поступившие в систему для дальнейшего распознавания и исследования:

$$Ag = (ag_1, ag_2, \dots, ag_m), \quad (11)$$

где m – общее количество антигенов.

Предложенный алгоритм состоит из следующих шагов:

1. Выбор антигена ag_j из популяции антигенов Ag .

2. Определение аффинностей $aff(ag_j; ab_i)$ для всей популяции антител Ab . Аффинность [4, 5] –

характеристика, количественно описывающая степень взаимодействия антигена и антитела:

$$\text{aff}(ag_j; ab_i) = (1 + d_{ij})^{-1}. \quad (12)$$

Здесь $\text{aff}(ag_j; ab_i)$ – аффинность между i -м антителом и j -м антигеном, а d_{ij} – Евклидово расстояние между ними.

3. К популяции антител Ab применяется алгоритм Зонге-Суня, с тем отличием, что избыточные пиксели не удаляются, а каждому антителу присваивается выживаемость.

4. Из популяции выбираются клетки с наихудшей выживаемостью и клонируются. Используется обратно-пропорциональное клонирование. При этом количество клонов для каждого антитела определяется в соответствии со следующим выражением:

$$Nc = n \cdot [1 - \text{aff}(ag_j; ab_i)], \quad (13)$$

где Nc – количество клонов, выделяемых антителу ab_j , а n – общее количество антител. После обратно-пропорционального клонирования получаем популяцию клонов C .

5. В каждый клон вносится пропорциональная направленная мутация - внесение изменений происходит после сравнения значений атрибутов клона (координаты, цвет, выживаемость клона) и ближайшего к нему по аффинности антигена. Если значение атрибутов клона превышает значение атрибутов антигена, происходит вычитание коэффициента из значения атрибута клона; если значение атрибута антигена превышает значение атрибута клона, то происходит добавление; в ситуации равенства значения атрибутов внесение изменений не происходит.

Коэффициент (уровень) мутации клонов зависит от аффинности их антитела-родителя с текущим антигеном:

$$\mu = \text{rand}[0; \text{aff}(ag_i; ab_j)]. \quad (14)$$

Таким образом, как результат, получаем популяцию новых клонов антител C' .

6. Для каждого антитела со слабой выживаемостью из популяции Ab вычисляется аффинность с каждой клеткой из C' как Евклидово расстояние.

7. В соответствии с полученными на шаге 4 аффинностями в популяции антител Ab проводится замена клеток с худшей аффинностью клонами [4, 5].

8. Проверка критерия останова, если критерий достигнут, то из популяции удаляются клетки с наихудшей выживаемостью [4, 5] и алгоритм останавливается, если нет, то – возвращаемся к шагу 1.

Предлагаемый метод позволяет повысить качество выполнения задачи скелетонизации символа за счет отсутствия этапа удаления избыточных клеток и обработки их параметров на основе модели клональной селекции.

Экспериментальные исследования

Для проведения экспериментальных исследований было создано приложение, демонстрирующее работу предложенного подхода для решения задачи скелетонизации символов, в качестве которых использованы буквы русского алфавита. Исследования проводились на монохромном изображении размером 100×100 точек. При реализации предложенного гибридного подхода применялось следующее кодирование антител: по одному байту на координаты, один бит на цвет ($\{0;1\}$ – белый или черный, соответственно), 3 бита для задания выживаемости клетки. Порог аффинности был принят равным 0.9. Выживаемость, устанавливаемая алгоритмом Зонге-Суня – $\{0;1\}$ (избыточна клетка или нет). Вносимая мутация лежит в пределах от 1 до 5.

В задаче скелетонизации в роли антигена выступает целевое изображение; в роли антител – выборка случайно сгенерированных точек. Для проверки работы приложения были проведены следующие эксперименты. Системе предлагалось построить скелет одного из выбранных символов на основе классического алгоритма Зонге-Суня и предложенного в работе подхода. На рис. 1–3 представлены результаты работы двух методов на примере символов «А», «С» и «Б».

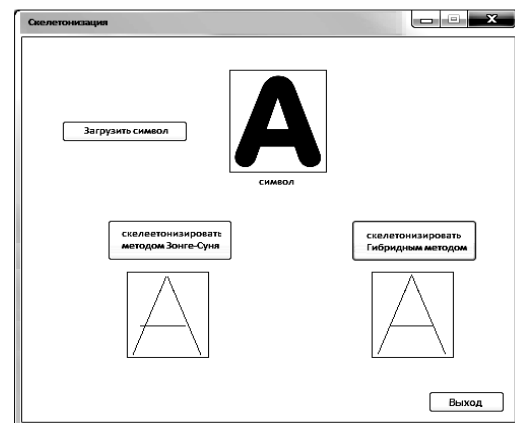


Рис. 1. Результаты скелетонизации символа «А»

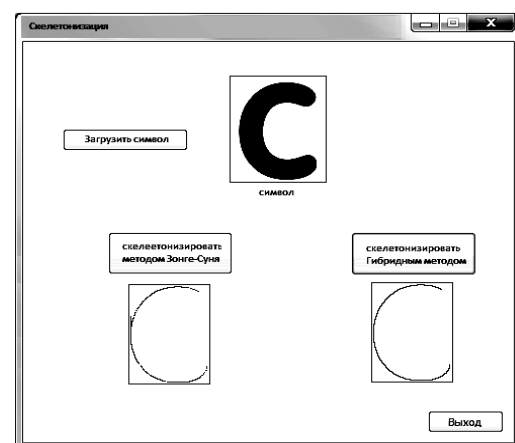


Рис. 2. Результаты скелетонизации символа «С»

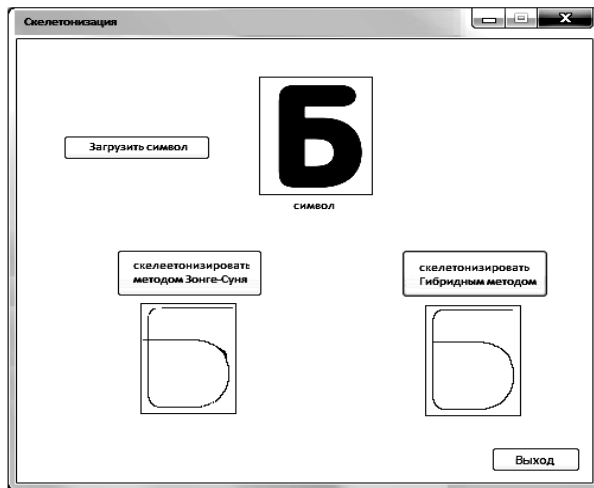


Рис. 3. Результаты скелетонизации символа «Б»

Используя алгоритм Зонге-Суэя для построения скелета символа, мы получаем остов с разрывами и избыточными точками, утолщающими кривую скелета.

Проведенные эксперименты позволяют сравнить результаты построения скелета символа на основе метода Зонге-Суэя и гибридного метода с использованием модели клональной селекции. Видно, что предлагаемый в работе гибридный подход позволяет избавиться от разрывов при построении скелета символа, и, как следствие, скелет не требует дальнейшей обработки.

Выводы

В работе рассмотрено решение актуальной задачи скелетонизации символов и предложен гибридный метод, который объединяет алгоритм скелетонизации Зонге-Суэя и иммунный подход, в частности, модель клональной селекции.

Данный метод позволяет получить более эффективное решение задачи предварительной обработки символов путем улучшения качества скелета

изображения, что дает возможность сразу выделить на нем признаки для распознавания без дополнительной обработки.

Результаты экспериментальных исследований показали, что применение модели клональной селекции к классическому алгоритму Зонге-Суэя дает возможность построения качественного скелета символа с отсутствием разрывов и утолщений остова символа.

Предложенный в работе гибридный метод может быть использован на последующих этапах распознавания символов.

Список литературы

1. Gonzales R.C. *Tsyfrovaya obrabotka izobrazheniy [Digital Image Processing]* / R.C. Gonzales, R.E. Woods. — Tekhnosfera, Moscow, Russia, 2002. — 1072 p.
2. Zhang T.Y. and Suen C.Y. *A fast parallel algorithm for thinning digital patterns* / T.Y. Zhang, C.Y. Suen // *Comm. ACM.* — 1984. — V. 27, No. 3. — P. 236 – 239.
3. Temnov K.A. *Raster monochrome images thinning algorithm* / K.A. Temnov, A.V. Kirillov, // *Informatsion-noizmeritelnie i upravlayu shchie sistemy.* — 2009. — No. 8. — P. 45 – 49.
4. *Искусственные иммунные системы и их применение* / Под ред. Д. Дасгупты; пер. с англ. А.А. Романюхи. — М.: ФИЗМАТЛИТ, 2006. — 344 с.
5. Cao Y. *An Immunogenetic Approach in Chemical Spectrum Recognition* / Y. Cao, D. Dasgupta, A. Ghose Ed., S. Tsutsui Ed. // *The book "Advances in Evolutionary Computation"*. — Springer-Verlag, 2003. — P. 897 – 914.
6. Барковская О.Ю. *Параллельная модификация алгоритма скелетонизации бинарного изображения* / О.Ю. Барковская, Н.Г. Аксак // *Восточно-Европейский Журнал передовых технологий.* — 2006. — 4/2 (22). — С. 65 – 68.

Поступила в редколлегию 26.03.2014

Рецензент: д-р техн. наук, проф. С.Г. Удовенко, Харьковский национальный университет радиоэлектроники, Харьков.

СКЕЛЕТОНІЗАЦІЯ СИМВОЛІВ ІЗ ВИКОРИСТАННЯМ МОДЕЛІ КЛОНАЛЬНОЇ СЕЛЕКЦІЇ

М.М. Кораблев, О.В. Легедіна, О.Г. Лебедев

У даній статті розглядається задача підготовки зображення для розпізнавання, а саме скелетонізація. Проведено аналіз існуючих алгоритмів рішення і запропонований гібридний метод, який заснований на використанні штучних імунних систем, а саме, моделі клональної селекції та алгоритмі скелетонізації Зонге-Суэя. Проведено експериментальні дослідження та визначено подальші напрямки досліджень, орієнтовані на вирішення задачі розпізнавання символів.

Ключові слова: штучна імунна система, розпізнавання символів, клональна селекція, афінність, антитіло, антиген, скелетонізація, алгоритм Зонге-Суэя.

SYMBOL SKELETONIZATION BASED ON THE MODEL CLONAL SELECTION

N.M. Korablev, A.V. Legedina, O.G. Lebedev

This article discusses the task of preparing images for recognition, namely skeletonization. The analysis of existing algorithms and solutions proposed hybrid algorithm, which is based on the use of artificial immune systems, namely, clonal selection model and skeletonization algorithm of Zhang-Suen. Experimental studies and identified further areas of research aimed at solving the problem of character recognition.

Keywords: artificial immune system, symbol recognition, clonal selection, affinity, antibody, antigen, skeletonization, algorithm of Zhang-Suen.