

РАСПОЗНАВАНИЕ И ВЫДЕЛЕНИЕ ГРАНИЦ ОДНОРОДНЫХ ОБЛАСТЕЙ ПРИ ПАРАЛЛЕЛЬНОЙ ОБРАБОТКЕ ГИСТОЛОГИЧЕСКИХ ИЗОБРАЖЕНИЙ

Актуальною в наш час є проблема розпізнавання та виділення границь однорідних областей, зокрема вона поширена в медицині. В даній роботі представляється алгоритм, що швидко та якісно розв'язує задачі такого роду, а також описується спосіб його реалізації при створенні програмного продукту. Також демонструються результати роботи програми при аналізі різних гістологічних зображень.

The problem of area recognition and border highlighting of homogeneous areas is very actual nowadays. In most of them the problem is to find and show borders of cells or cell organoids quickly and clearly for further analysis. An algorithm to solve such problems is proposed in this article. The way of implementing the algorithm in software demonstrated. Results of this implementation are also shown.

1. Постановка проблемы

Целью любых медицинских исследований чаще всего является принятие решения в отношении предоставления помощи пациенту. От правильности принимаемого решения зачастую может зависеть жизнь человека. Поэтому очень важно, чтобы основания для выбора того или иного способа воздействия были наиболее точны и верны. В частности при гистологическом анализе огромное значение имеет правильное, четкое и наиболее быстрое получение результатов для последующего анализа и принятия необходимого решения. Медики анализируют количество клеток, попавших на предметное стекло, или же их размер, но начало анализа непременно состоит в выделении этих самых клеток, а иногда и их частей. Вот тут и возникает проблема наиболее верного и качественного выделения границ областей гистологических изображений.

2. Анализ последних исследований и публикаций

При исследовании данного вопроса были проанализированы такие методы выделения областей:

- Методы с использованием гистограмм;
- Методы разрастания областей;
- Методы водораздела;
- Сегментация с помощью модели
- Методы разрезания графа;
- Выделение границ.

Методы с использованием гистограмм подразумевают построение гистограмм по всем пикселям изображений для дальнейшего обнаружения минимумов и максимумов яркости, которые и принимаются за центры кластеров.

Вокруг этих центров и происходит дальнейшее формирование (выделение) однотипных зон. Эти методы не подходят для наших изображений, так как зачастую последние недостаточно яркие и контрастные. Методы разрастания базируются на выборе «семян» – точек в областях, которые мы хотим выделить. Из этих семян будут произрастать области, но задавать их нужно вручную для каждого конкретного рисунка, что делает данные методы слишком неудобными. Также они не исключают ошибок, связанных с человеческим фактором, что недопустимо для области наших исследований. Методы водораздела работают с градиентом яркости, представляя границы областей возвышениями, с которых вода стекает к центрам зон и, собираясь там, формирует наши зоны. Проблема данных методов в том, что они не могут обеспечить нам достаточного качества результатов. Сегментация с помощью модели подразумевает наперед определенную структуру изображений и основывается на предположении, что выделяемые области имеют некоторую наперед известную форму, что никак не подходит для поставленной задачи, ведь структура и форма клеток бывает произвольна. Методы разреза графа считаются достаточно универсальными и имеют хорошую производительность. Они основаны на том, что изображение представляется как взвешенный неориентированный граф. В чистом виде такие методы слишком специфичны, но мы воспользуемся данным представлением графической информации. Методы выделения границ являются наиболее подходящими, так как базируются на

утверждении, что на границе областей существует значительный перепад яркости. Из выбранных групп алгоритмов можно выделить две, которые лягут в основу выводимого нами алгоритма. Они называются алгоритмами «Краб» (от КРАТчайший Путь) и «Форель» (от первых букв ФОРмальный АЛгоритм)[1,3]. На основании проведенного анализа в следующем разделе будет представлен выведенный для данной задачи алгоритм.

3. Формулирование цели статьи

Целью данного исследования является выведение адекватного алгоритма для нахождения и выделения границ однотипных областей, который показывал бы хорошие результаты применимо к разным гистологическим изображениям, а также мог быть реализован в виде программного обеспечения с целью частичной автоматизации подобных исследований. Целью конкретной статьи является демонстрация проведенных исследований и полученных результатов.

4. Описание метода

После проведенного анализа мы остановились на использовании собственного алгоритма, который может быть представлен как производная алгоритмов «Краба» и «Фореля»[1,3]. Также его идея похожа на алгоритм Собеля решения подобных задач. Его суть заключается в том, что каждая точка изображения (пиксель) представляется как вершина графа, которая в пространстве имеет 5 координат, 3 из которых отвечают за составляющие цвета (Red, Green, Blue), а 2 оставшихся за положение в пространстве (x и y). В связи с тем, что работа с пятимерным пространством весьма затруднительна, первым делом мы переводим изображение в оттенки серого согласно известной формуле:

$$K = 0,11 * B + 0,3 * R + 0,59 * G$$

где K – оттенок серого, приобретаемый точкой; R, G, B – значения красной, зеленой и синей составляющих цвета. Для преобразования изображения программно значение K заносится в компоненты R, G и B , обеспечивая серый цвет.

После этого необходимо как-то выделить границы областей. Для этого сначала высчитываем разностную картину. Проводится следующее преобразование:

$$z_j = z_j - z_{j+1}, \quad j = 1, \dots, n - 1$$

Согласно вышеуказанной формуле, проходя по пикселям всего изображения, в качестве цвета текущей точки берется значение разности цветов текущего и следующего пикселей. Получается оттенок серого, который будет темнее, чем больше перепад контрастности (как на границах областей). После этого, мы получаем картину, где на белом фоне уже различимы слабо проявленные границы и шум.

На следующем этапе необходимо сделать границы более заметными. Для этого мы проводим бинаризацию изображения: проходим по всем пикселям изображения и меняем их цвет на черный или белый. Эта операция проводится относительно определенного коэффициента, который удобно подбирать экспериментально для каждой группы изображений. Пиксель перекрашивается в черный если оттенок серого больше коэффициента, иначе в белый цвет. Его выбор зависит от четкости изображений и контрастности на границах областей. Наилучшие результаты в наших исследованиях показал коэффициент в пределах 243-249.

На предыдущем шаге было получено черно-белое изображение, где черные пиксели являются границами областей и шумом. Теперь, проходя по пикселям изображения, мы можем считать граничные точки. Простое считывание подряд не удовлетворяет поставленной задаче, так как порождает проблему определения принадлежности граничной точки той или иной области, потому был выбран метод «ближайшего соседа»[2]. Сначала надо найти любую первую черную точку, относительно которой и будем искать соседние. Так как пиксели лежат упорядочено, поиск соседней черной точки будем проводить в квадрате вокруг неё. Границы квадрата будем расширять на 1 если следующая черная точка не была найдена (схема на рис. 1).

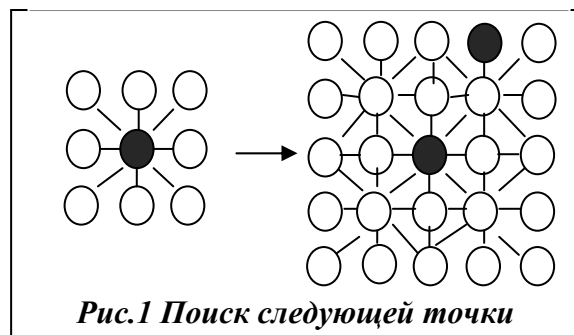


Рис.1 Поиск следующей точки

Находя пиксели таким образом, получаем последовательность точек границ зон. Чтобы разделить их по принадлежности разным зонам

(или добавит каждой точке дополнительный атрибут, показывающий номер зоны, которой данная точка принадлежит), нужно разбить эту последовательность на отдельные области. Критерием разбиения на области может быть расстояние между соседними найденными точками (превышает ли оно определенную константу, определяемую в зависимости от типа изображения).

В списке найденных граничных точек содержится также и шум, который необходимо отфильтровать. Если коэффициент, по которому различаются разные зоны, подобран верно, шум отсеивается следующим образом. Предполагается, что правильно выделенные границы областей не могут содержать менее некоторого числа точек (эта величина подбирается в зависимости от величины объектов, которые следует распознать и величины расстояния между ними). Проходя по всем выделенным точкам, считаем количество граничных пикселей каждой зоны и если их количество менее заданного, область удаляется из рассмотрения так как признается шумом. Этот метод очистки изображения позволяет улучшить точность распознавания границ и исключить из дальнейшего анализа лишней информации.

На основе вышеизложенного алгоритма был создан программный продукт для распознавания границ однотипных областей изображений. Его испытание было проведено на гистологических изображениях, хотя алгоритм может использоваться и для других предметных областей.

Покажем далее результаты применения нашего алгоритма к реальным гистологическим изображениям. На рис. 2 показано начальное тестовое изображение, границы областей которого необходимо распознать, а на рис. 3 – то же изображение, но с распознанными границами.

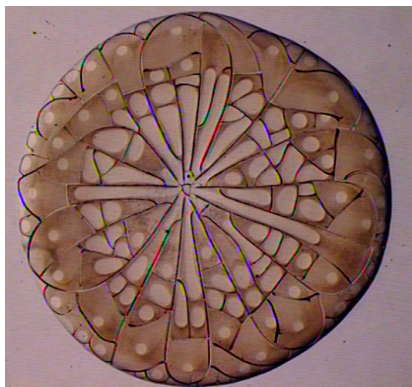


Рис. 2 Клетка крови до выделения границ

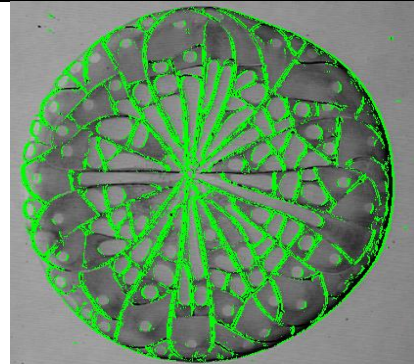


Рис. 3 Клетка крови с распознанными границами (зеленый цвет)

Также, алгоритм применим и для распознавания границ нескольких разных клеток (смотри рис. 4 и 5).

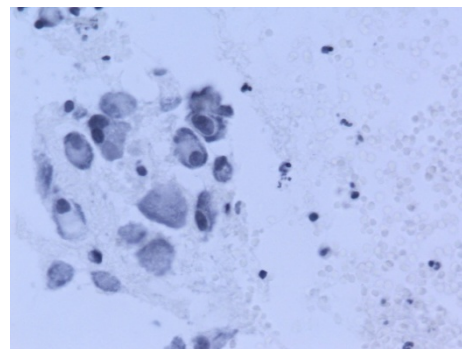


Рис. 4 Онкологическое начальное изображение

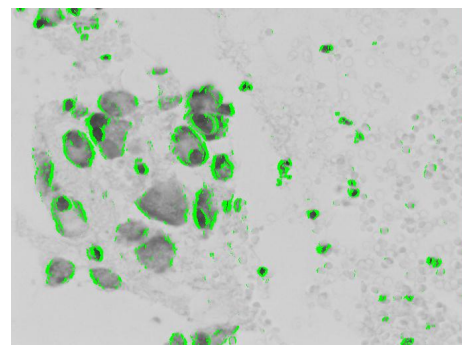


Рис. 5 Изображение (рис. 4) с распознанными границами (зеленый цвет)

5. Анализ полученных результатов

Для оценивания качества полученных результатов был выбран метод экспертного оценивания. Была сформирована выборка тестовых изображений разного характера, аналогичных приведенным выше. Она анализировалась полученной программой, а результаты анализа вместе с исходными изображениями показывались экспертам, которые выставляли каждой паре изображений балл от 0 до 10. Потом данные по всей выборке и по всем экспертам усреднялись, и полученный средний балл принимался как оценка всей системы. Результаты данного анализа показаны на Рис. 6.



Рис. 6 График оценки качества полученного программного продукта

Как видно на рис. 6, оценки являются достаточно высокими (среднее значение 9,02 из возможных 10). Из этого можно сделать вывод, что программа достаточно эффективно справляется с поставленной задачей.

6. Выводы

В результате проделанной работы был создан алгоритм, позволяющий эффективно выделять границы однотипных областей изобра-

жений. Он является вполне универсальным и может быть применен для различных предметных областей. В рамках этого исследования был создан программный продукт, реализующий приведенный выше алгоритм. Тестирование программы проводилось на гистологических изображениях и были получены удовлетворительные результаты.

Как видим на рисунках 2 и 3, алгоритм недостаточно хорошо распознает горизонтальные линии. Для исправления этого недостатка при нахождении разностной характеристики достаточно добавить еще один проход по пикселям: определять разность цветов между вертикально расположенными пикселями.

Работа может стать основой для принятия решений в различных сферах человеческой жизни, в частности, в медицинских экспертных системах. Разработанный программный продукт может быть использован при создании экспертной онлайн – системы как её составляющая или как независимое программное обеспечение при исследовании графической информации.

Список литературы

1. Ципилева Т.А. Геоинформационные системы: Учебное пособие / Ципилева Т.А.– Томск: (Томский межвузовский центр дистанционного образования, 2004).
2. Загоруйко Н.Г. Алгоритмы обнаружения эмпирических закономерностей. / Загоруйко Н.Г., Елкина В.Н., Лбов Г.С Новосибирск: (Наука.-1985)
3. Вапник В.Н. Теория распознавания образов / Вапник В.Н., Червоненкис А.Я. Москва: (Наука, 1974)