

ПРОЦЕДУРА СЕГМЕНТАЦИИ МЕДИЦИНСКИХ ИЗОБРАЖЕНИЙ, ПОЛУЧЕННЫХ ПОД МИКРОСКОПОМ

Оробинский П.А.

Харьковский национальный университет радиоэлектроники, УКРАИНА

АННОТАЦИЯ. В данной работе представлены теоретические и практические методы, позволяющие производить подсчет различных однотипных объектов на полутоновых изображениях, полученных под микроскопом. В частности рассмотрены проблемы и способы их решения по нормализации, грубой сегментации, поиска стыковых точек и границ объектов на конгломератах.

Введение. Актуальным и обширным полем для научных исследований являются решение задачи нормализации и сегментации на полутоновых изображениях, полученных под микроскопом. Сложность решения таких задач обусловлена неоднозначность использования предварительного этапа обработки входного изображения – фильтрации для удаления шума, изменения контраста для повышения качества восприятия микрообразов [1, 2].

Цель работы. Целью данного исследования является разработка максимально эффективного алгоритма для нормализации и последующей сегментации изображения. Под критерием эффективности алгоритма будем понимать минимально возможное количество итераций для достижения лучшего результата с некой заданной вероятностью $P(a)$. В качестве предметной области выбрана задача обнаружения количества $E. coli$ на цифровых снимках, полученных под микроскопом.

Основная часть работы. Для решения поставленной задачи был предложен следующий алгоритм:

1. Нормализация входного изображения. Нормализатором в данном случае служит линейный оператор, отображающий исходное изображение в нормализованное, полученный из отношения математического ожидания входного изображения к математическому ожиданию эталонного. Под эталонным изображением будем понимать такое, что множества пикселей, принадлежащих фону, оболочке объекта и телу не пересекаются, т.е.

$$A = A_1 \cup A_2 \cup A_3, A_1 \cap A_3 = \emptyset, A_1 \cap A_2 = \emptyset, A_2 \cap A_3 = \emptyset, \quad (1)$$

где A – множество всех пикселей входного изображения, A_1 – пиксели, принадлежащие фону, A_2 – оболочке объекта, A_3 – телу объекта.

2. Фильтрация на основе статистических данных. В качестве первичного грубого фильтра предлагается считать единичным объектом такое множество, которое удовлетворяет следующим условиям:

$$(S_{min} \leq S \leq S_{max}) \& (\rho_{min} \leq \rho \leq \rho_{max}) \& (k_{min} \leq k \leq k_{max}), \quad (2)$$

где S_{min} и S_{max} , ρ_{min} и ρ_{max} , k_{min} и k_{max} – минимальные и максимальные значения площади, яркости, отношения площади оболочки к площади тела объектов соответственно, полученные на основании проведенных статистических исследований (табл. 1).

Таблица 1. Минимальные и максимальные значения площади, длины окружности, яркости тела и оболочки нормализованного объекта, отношение площади оболочки к телу объекта в серии из 100 испытаний.

Параметр	Min	Max
Площадь(пиксели)	492	896
Отношение площади оболочки к телу объекта	0,34	0,43
Длина окружности объекта (пиксели)	75	140
Яркость тела объекта(нормализованного объекта)	60	75
Яркость оболочки объекта(нормализованного объекта)	76	104

3. Сегментация конгломератов объектов, где под конгломератом понимаются склеившиеся объекты, т.е. площадь значительно превышает – $S \gg S_{max}$.

4. Подсчет объектов, полученных в результате фильтрации входного изображения и сегментации конгломератов.

Для сегментации конгломерата были определены точки стыка (ТС) объектов. Точкой стыка (x, y) является точка локального экстремума (ТЛЭ), при этом функция на отрезке $(x - m, x)$ не возрастает или не убывает и, соответственно, на отрезке $(x, x + m)$ не убывает или не возрастает по dy , либо аналогично по dx для $(y - m, y)$ и $(y, y + m)$, где m – некая эмпирическая константа (примерно 5% от средней длины окружности объекта – 7 пикселей). Под границей будем понимать последовательность 8-связных пикселей [3], не лежащих на одном уровне дерева (с корнем в ТС), соединяющую ТС с ближайшим пикселем, принадлежащим фону. Оптимальным на каждом уровне узлом будем считать такой, математическое ожидание (МО) яркостей пикселей 5×5 окрестности которого стремиться к МО яркостей границы эталонного объекта.

Задача поиска оптимальной границы сводится к задаче поиска оптимального пути на О-графе. В связи с тем, что известные методы решения данной задачи ресурсоемки, то было предложено разбить граф на подграфы, при этом поиск оптимального пути рассматривать, как поиск такового с некоторой заданной вероятностью $P(a)$. В результате для минимизации ресурсов требуется разрешить задачу оптимизации (ЗО) – определить такое разбиение графа на подграфы, чтобы путь, найденный за некоторое количество итераций, был оптимален с вероятностью $P(a)$, а количество итераций для его нахождения было минимально. Входными параметрами для этой задачи является требуемая вероятность оптимальности пути, вероятность помехи, максимальная длина пути (глубина дерева). Для решения данной ЗО была построена база знаний, которая определяет оптимальные глубины подграфов для заданных вероятности помехи и требуемой вероятности оптимальности пути. В частности, для рассматриваемой задачи вероятности помехи не превышала 0.1, а длина границы не превышала 60 пикселей, требуемая вероятность оптимальности пути – 0,99. Так, заданную вероятность оптимального пути можно достичь за 10 повторений при разбиении исходного графа на 10 частей. Общее количество итераций не превысит 58 563 (алгоритм Дейкстры для графа с 60 уровнями составит $n * m = 37\ 594\ 800$ итераций, где $n = 3\ 600$ – количество вершин, $m = 10\ 443$ – количество ребер).

Выводы. Предложенный теоретический подход был успешно применен на практике (реализовано программное средство на C++) – в 100% исследуемых изображений расчетное количество объектов совпало с реальным. Главным достоинством предложенного подхода является скорость (относительно небольшое количество итераций), которая в некоторых случаях превышает классические алгоритмы [4] (Дейкстры, Форда-Белмана, Флойда-Уоршелла) в сотни раз. К основным недостаткам следует отнести сложность предварительных расчетов – затраты на построение необходимой базы знаний. Еще одним недостатком можно считать, что данный метод не дает 100% вероятности нахождения оптимального пути за проделанное количество шагов. Однако, учитывая, что задача состоит в том, чтобы посчитать именно количество объектов на изображении, то некой погрешностью границ можно пренебречь. Открытыми остаются вопросы применимости данного метода для разделения объектов с границей, много превышающей исследуемые 60 пикселей, и для изображений с высоким уровнем шума.

СПИСОК ИСПОЛЬЗОВАННЫХ ИСТОЧНИКОВ

1. Lyashenko, V. V., Babker, A. M. A. A., & Kobylin, O. A. (2016). The methodology of wavelet analysis as a tool for cytology preparations image processing. *Cukurova Medical Journal*, 41(3), 453-463.
2. Lyashenko, V., Matarneh, R., Kobylin, O., & Putyatin, Y. (2016). Contour detection and allocation for cytological images using Wavelet analysis methodology. *International Journal of Advance Research in Computer Science and Management Studies*, 4(1), 85-94.
3. Гонсалес Р., Вудс Р., Эддинс С. Цифровая обработка изображения в среде MATLAB. – М.: Техносфера, 2006. – 616 с.
4. Томас Х. Кормен, Чарльз И. Лейзерсон, Рональд Л. Ривест, Клиффорд Штайн. Алгоритмы: построение и анализ - Introduction to Algorithms. — 2-е изд. — М.: Вильямс, 2006. — 1296 с.