

**МЕТОД СЕГМЕНТАЦИИ ЛЕЙКОЦИТОВ НА ИЗОБРАЖЕНИЯХ МАЗКОВ КРОВИ**

© 2012 г. А. В. Дырнаев, студент

Санкт-Петербургский национальный исследовательский университет информационных технологий, механики и оптики, Санкт-Петербург

E-mail: dirnalex@yandex.ru

В статье изложены особенности метода сегментации микроскопических изображений препаратов крови с выделением основных типов лейкоцитов и их подсчетом. Сегментация производится на основе порогового разделения изображения мазка крови по яркости, тону и насыщенности с использованием морфологической обработки бинарных масок, расчета параметров объектов, таких как коэффициент заполнения, эксцентриситет и эквивалентный диаметр, расчета текстурных характеристик объектов. Метод устойчив к присутствию шумов на изображении, а также к небольшим изменениям цветовых характеристик клеток.

*Ключевые слова:* подсчет количества клеток, цифровая микроскопия, сегментация изображений.

Коды OCIS: 150.1135

*Поступила в редакцию 29.05.2012***Введение**

При анализе мазков крови важнейшим параметром, который требует расчета, является лейкоцитарная формула [1]. Она представляет собой соотношение различных типов белых клеток крови (лейкоцитов). Среди лейкоцитов различают гранулоциты и агранулоциты. К агранулоцитам относят лимфоциты и моноциты. К гранулоцитам относят нейтрофилы, разделяющиеся на сегментоядерные и палочкоядерные, базофилы и эозинофилы. Лейкоцитарная формула позволяет следить за ходом патологического процесса, прогнозировать исход болезни, судить о наличии осложнений. По отклонению концентрации тех или иных клеток от нормы можно судить об изменениях в органах, а также о функционировании кроветворной системы организма. Увеличение количества лейкоцитов (лейкоцитоз) возникает при острых воспалительных процессах и многих инфекционных заболеваниях, злокачественных новообразованиях, травмах тканей, инфаркте миокарда, при беременности, после родов и больших физических нагрузок (физиологический лейкоцитоз). К снижению числа лейкоцитов (лейкопения) приводят воздействие ионизирующего излучения, брюшной тиф, вирусные заболевания, анафилактический шок, коллаптозы, влияние некоторых

лекарственных препаратов, острые лейкозы, и многие другие факторы. Вместе с подсчетом эритроцитов на изображениях мазков крови [2], расчет лейкоцитарной формулы, а также расчет количества лейкоцитов составляют основную часть общего анализа крови.

Задачу сегментации лейкоцитов решают различными способами: на основе морфологического критерия [3], с использованием информации о цвете [4], с помощью нейронной сети [5], на основе статистической информации по различным критериям [6], при анализе контурного изображения [7, 8]. Однако перечисленные методы обладают рядом недостатков, таких как недостаточно высокая вероятность правильной классификации, не превышающая 90% [6], требование обширной базы изображений клеток для обучения [5], требование к низкой зашумленности изображения [7, 8] или к окраске изображения [4]. В статье предложен метод, основанный на комплексном применении различных методов, что позволяет повысить достоверность сегментации, а также понизить требования к тону изображения и к его качеству.

**Типы сегментируемых клеток**

Предлагается производить сегментацию лейкоцитов, учитывая особенности окраски,

текстуры, структуры и размеров каждого из типов. Различают следующие типы лейкоцитов: лимфоциты, моноциты, нейтрофилы, эозинофилы и базофилы.

**Лимфоциты** участвуют в распознавании антигенов. В лейкоцитарной формуле в норме их содержание должно быть 19–37%. При окраске по Романовскому–Гимзе [9] цитоплазма лимфоцитов окрашивается в сине-голубой цвет, их ядра – в цвета от интенсивно пурпурного до фиолетового. Ядро – круглое, гиперхромное (сильно окрашенное). Цитоплазма составляет лишь узкий ободок вокруг ядра, базофильна и не содержит гранул. Размер лимфоцитов колеблется от 4,5 мкм – у малых форм, 7–10 мкм – у средних и от 10 мкм – у больших.

**Моноциты** удаляют из организма отмирающие клетки, остатки разрушенных клеток, денатурированный белок, бактерии и комплексы антиген-антитело. В лейкоцитарной формуле их содержание 3–11%. Цитоплазма моноцитов окрашивается в мутный голубовато-серый цвет, их ядра – в более светлый пурпурно-красный, присутствует и азурофильная зернистость. Ядро – бобовидное или подковообразное, светлое. Цитоплазма составляет заметную часть клетки, менее базофильна, чем у лимфоцитов, и возле ядра может содержать несколько гранул. Моноциты являются самыми крупными клетками крови, их размер в капле свежей крови оставляет до 12 мкм, в мазке до 18–20 мкм.

**Нейтрофилы** обладают, в основном, бактерицидной и дезинтоксикационной функциями. В норме представлены тремя группами: юные, палочкоядерные и сегментоядерные. Соотношение в лейкоцитарной формуле для каждой группы соответственно: 1–5%, 1–5% и 40–68%. Гранулы нейтрофилов окрашиваются в цвета от пурпурного до фиолетового. Так как нейтрофилы являются гранулоцитами, они имеют два характерных признака: специфическую зернистость в цитоплазме и сегментированное ядро. Ядро нейтрофилов состоит из нескольких (обычно 3–4) связанных друг с другом сегментов, а в цитоплазме присутствует трудно различимая мелкая зернистость, обусловленная наличием гранул фиолетово-розового цвета. У палочкоядерных нейтрофилов ядро имеет форму изогнутой палочки, а у юных – бобовидную форму. Нейтрофильные гранулы по размеру равны 0,1–0,3 мкм, в то время как диаметр клетки в среднем равен 8 мкм.

**Эозинофилы** обладают фагоцитарными свойствами, но это свойство используют прежде всего для участия в аллергическом процессе. У взрослых людей в норме в крови содержится 0,5–5,0% данных клеток. У эозинофилов присутствуют палочкоядерные и юные формы, но из-за малого содержания их в крови они не включены в лейкоцитарную формулу. Гранулы эозинофилов окрашиваются по Романовскому–Гимзе в оранжево-розовый цвет. Ядро эозинофильного гранулоцита обычно состоит из двух сегментов. В цитоплазме преобладают оксифильные гранулы, окрашенные эозином в розовый цвет. Эозинофильные гранулы крупные (0,5–1,5 мкм), овальной или полигональной формы. Размер эозинофилов колеблется в пределах 9–10 мкм.

**Базофилы** участвуют в воспалительных и аллергических процессах в организме. В норме в крови содержится до 1% данного типа клеток. Гранулы крупные (0,5–1,2 мкм), заполняют почти всю цитоплазму базофила. Цитоплазма содержит большое количество крупных базофильных гранул фиолетово-вишневого цвета, сквозь которые просматривается ядро. У базофилов ядро имеет слабодольчатую структуру, но разглядеть последнюю не всегда удается. Гранулы базофилов окрашиваются в интенсивный сине-фиолетовый цвет. Размер базофилов в среднем около 9 мкм.

## Предлагаемый метод

Так как фон на изображениях мазков крови всегда светлее объектов, первым этапом предлагаемого метода является сегментация изображения с определением фоновой составляющей по гистограмме яркости изображения мазка крови. Порог выбирается с помощью метода Отсу [10], основанного на минимизации внутрикластерных расстояний. Полученная бинарная маска инвертируется так, чтобы ее единичные значения соответствовали фону.

Далее изображение мазка крови переводится в цветное координатное пространство HSV и строится гистограмма распределения значений компонент тона с учетом маски фона и “не фона”. Иными словами, для построения гистограммы берут только те элементы изображения, которые не принадлежат фону. На гистограмме присутствуют два максимума, соответствующие белым кровяным клеткам и остальным клеткам. Поиск максимума проводится после низкочастотной Фурье фильтрации

гистограммы изображения. Фурье фильтрация представляет собой фильтрацию в частотной области. Маска фильтра, применяемая в частотной области, приравнивает все значения, больше частоты среза фильтра, к нулю. Частота среза увеличивается до тех пор, пока на гистограмме фильтрованного изображения не появятся два максимума. Поиск максимумов производится с помощью дифференцирования гистограммы фильтрованного изображения. В местах смены знака производной гистограммы фильтрованного изображения находятся максимумы или минимумы гистограммы. Как только два максимума будут найдены на гистограмме, минимум между ними принимается за искомым, и соответствующее ему значение тона будет принято за пороговое. По пороговому значению изображение снова сегментируется с учетом маски фона и “не фона”. Результатом преобразований является бинарная маска, определяющая местоположение лейкоцитов на изображении мазка. После применения данной маски удаляются малые ложносегментированные области. Производится морфологическая операция закрытия и заполнения дыр в объектах.

На полученной маске объекты разделяются и нумеруются путем селекции связных компонент [11]. Селекция связных компонент представляет собой присвоение каждой области, окруженной фоном, уникальной метки. Таким образом, появляется возможность работать исключительно с изображениями белых кровяных клеток. В связи с тем, что моноциты значительно крупнее лейкоцитов, эти клетки могут быть определены благодаря своему большому размеру.

Затем на оставшихся клетках кластеризацией множества значений яркости и насыщенности методом  $k$ -средних [12, 13] выделяется ядро клетки. Далее рассчитывается коэффициент заполнения как отношение площади объекта к площади выпуклой фигуры, в которую можно вписать объект. Поскольку ядро лимфоцита выпуклое и занимает большую часть клетки, по его коэффициенту выпуклости и соотношению площади ядра и цитоплазмы можно определить клетку-лимфоцит.

Если ядро разделено на две части, то можно сказать, что это эозинофил. Базофил определяется с помощью текстурной сегментации на основе измерения среднего значения параметра однородности, рассчитанного на основе матрицы смежности. И, наконец, сегментоядерный

нейтрофил отличается от палочкоядерного и от юного по форме ядра. Вычисляя эксцентриситет ядра можно сказать, вытянуто оно, как в случае с палочкоядерным нейтрофилом, или наоборот, эксцентриситет выпуклой фигуры, в которую вписывается ядро, близок к нулю, что указывает на то, что исследуемая клетка принадлежит к сегментоядерным нейтрофилам.

## Эксперимент

Для экспериментальной проверки метода были взяты изображения, содержащие лейкоциты различных видов. Пример тестового изображения приведен на рис. 1. Было выбрано одно изображение с палочкоядерным нейтрофилом, два изображения с базофилами, два изображения с эозинофилами, три изображения с лейкоцитами, два с моноцитами, и два с сегментоядерными нейтрофилами. С помощью метода Оцу был удален фон из тестовых изображений, а затем по цветовой характеристике были сегментированы лейкоциты на изображениях.

В результате сегментации было получено бинарное изображение со значением равным единице в тех местах, где предположительно находятся лейкоциты. Из бинарного изображения были удалены объекты, площадь которых меньше, чем площадь наибольшего объекта в семь раз. Такой порог был определен экспериментальным путем, так как при меньшем пороге наблюдалось присутствие на бинарном изображении посторонних объектов, а при большем наблюдалось исчезновение объектов, соответствующих искомым клеткам.

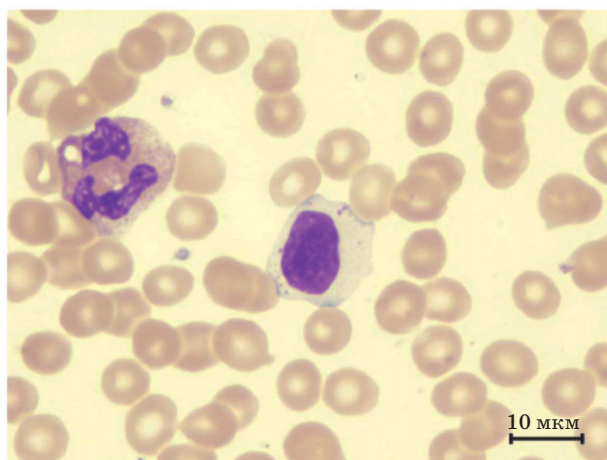
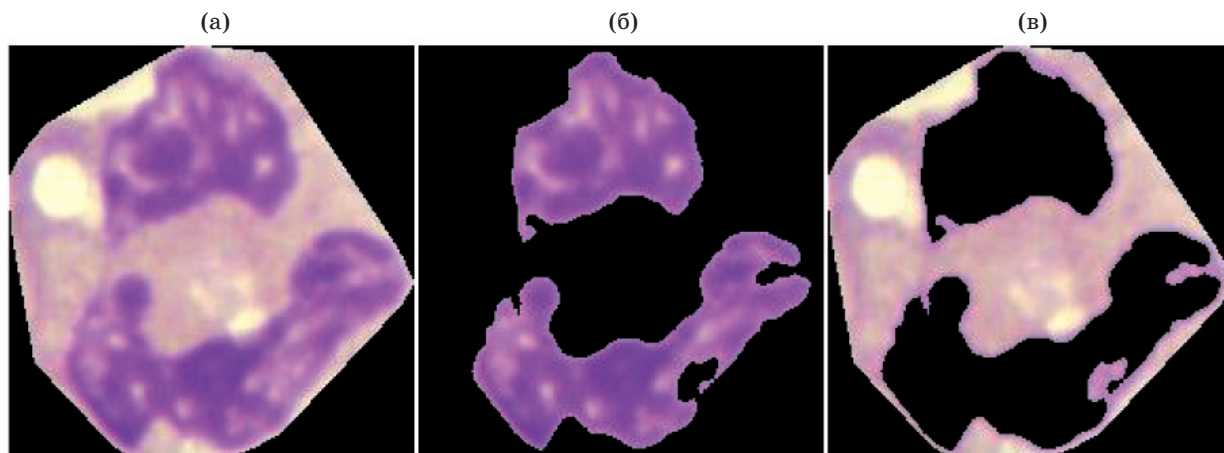


Рис. 1. Пример тестового изображения.



**Рис. 2.** Пример сегментации лейкоцита и его ядра. а – выделенный лейкоцит, б – выделенное ядро лейкоцита, в – выделенная цитоплазма лейкоцита.

Затем была проведена операция закрытия изображения со структурным элементом размером  $5 \times 5$ . На бинарном изображении были заполнены единичными значениями все “дыры” в объектах.

Полученная бинарная маска была разделена на связные компоненты и для каждой компоненты рассчитаны следующие параметры: изображение ограничивающего многоугольника, координаты левого верхнего угла, ширина и высота ограничивающего прямоугольника, эквивалентный диаметр (диаметр круга, имеющего такую же площадь). Изображение выделенного лейкоцита приведено на рис. 2а. Если масштаб изображения неизвестен, предлагается определить его по среднему размеру эритроцитов. Маска лейкоцитов и других клеток крови инвертируется, из полученного бинарного изображения удаляются все элементы размером более 500, дыры в объектах заполняются единичными значениями. Из полученного изображения выбираются только те объекты, эксцентриситет которых менее 0,5, а коэффициент заполнения более 0,9. Средний эквивалентный диаметр полученных объектов принимается за средний размер эритроцитов в элементах изображения. Поскольку средний диаметр эритроцита равен 7,5 мкм, масштаб изображения будет равен отношению среднего диаметра в элементах изображения к среднему диаметру в микрометрах.

Для каждой найденной клетки производится кластеризация на два уровня по насыщенности и яркости, тем самым определяется ядро клетки и цитоплазма. Пример разделе-

ния клетки на цитоплазму и ядро приведен на рисунках 2б, в. Рассчитывается среднее значение насыщенности и яркости для ядра и цитоплазмы.

Значения количества ядер, коэффициента заполнения ядра, эксцентриситета ядра, размера клетки, расстояния от центра масс ядра до центра масс цитоплазмы, разности между средним значением яркости ядра и цитоплазмы и разности между средним значением насыщенности ядра и цитоплазмы вычисляются для каждой найденной клетки, а затем эти параметры сравниваются с характерными значениями для каждого типа клеток. Полученные значения этих параметров для тестовых изображений представлены в табл. 1. В результате каждый лейкоцит отнесен к определенному типу. Тип лейкоцита, определенный предложенным методом, и его истинный тип указаны в табл. 2.

Результат сегментации не зависел ни от тона изображения, ни от степени зашумленности. На выборке из 14 клеток было правильно классифицировано 13 клеток.

## Выводы

На выборке из 12-ти изображений, содержащих 14 лейкоцитов всех возможных типов, лишь одна клетка была определена неправильно, так как ее цитоплазма была очень слабо различима. Таким образом, метод пригоден для использования в задачах анализа образцов крови при различных лабораторных условиях. Метод применим для анализа не только клеток крови человека, но может быть использо-

**Таблица 1.** Значения рассчитываемых параметров для обнаруженных лейкоцитов

Номер клетки	Количество ядер	Коэффициент заполнения ядра	Эксцентриситет ядра, мкм	Размер клетки, мкм	Расстояние от центра масс ядра до центра масс цитоплазмы, мкм	Разница между средним значением насыщенности ядра и цитоплазмы	Разница между средним значением яркости ядра и цитоплазмы
1	1	0,647	0,899	16,285	9,182	52,739	46,359
2	1	0,746	0,673	16,659	7,054	20,377	29,811
3	1	0,461	0,830	18,209	8,309	78,678	47,136
4	2	не рассчитывается	не рассчитывается	не рассчитывается	не рассчитывается	22,855	29,077
5	2					102,007	47,420
6	1	0,592	0,508	10,271	5,297	14,120	12,197
7	2	не рассчитывается	не рассчитывается	не рассчитывается	не рассчитывается	65,263	30,509
8	1	0,963	0,730	12,027	1,586	64,905	21,616
9	1	0,971	0,439	12,600	2,394	103,785	37,097
10	1	0,919	0,751	17,130	7,392	44,919	27,817
11	1	0,882	0,702	17,882	3,180	48,675	25,734
12	1	0,892	0,665	22,251	2,490	118,227	48,943
13	1	0,742	0,747	14,168	1,715	98,528	62,801
14	4	не рассчитывается	не рассчитывается	не рассчитывается	не рассчитывается	36,559	41,150

ван и в ветеринарии для анализа клеток крови животных. В этом случае требуется изменение параметров метода, таких как средний размер клеток или характеристики ядра основных типов клеток. Возможно применение предложенного метода при сегментации клеток других типов при условии наличия у них характерной окраски и ядра.

Достоинство предложенного метода состоит в автоматическом определении масштаба изображения по среднему размеру эритроцитов. Программная реализация метода, выполненная в среде программирования MATLAB, показала высокое быстродействие при обработке изображений мазков крови. В дальнейшем планируется модификация метода для сегментации слипшихся клеток, а также для повышения эффективности сегментации в условиях одинакового тона окраски эритроцитов и лейкоцитов.

Работа выполнена при финансовой поддержке Министерства образования и науки Российской Федерации.

**Таблица 2.** Результат сегментации лейкоцитов

Номер клетки	К какому типу отнесена клетка	К какому типу клетка относится на самом деле
1	Палочкоядерный нейтрофил	Палочкоядерный нейтрофил
2	Базофил	Базофил
3	Базофил	Базофил
4	Эозинофил	Эозинофил
5	Эозинофил	Эозинофил
6	Лимфоцит	Базофил
7	Эозинофил	Эозинофил
8	Лимфоцит	Лимфоцит
9	Лимфоцит	Лимфоцит
10	Моноцит	Моноцит
11	Моноцит	Моноцит
12	Моноцит	Моноцит
13	Сегментоядерный нейтрофил	Сегментоядерный нейтрофил
14	Сегментоядерный нейтрофил	Сегментоядерный нейтрофил

\* \* \* \* \*

## ЛИТЕРАТУРА

1. *Афанасьев Ю.И., Юрина Н.А., Котовский Е.Ф.* Гистология, цитология и эмбриология // М.: Медицина. 2002.
2. *Дырнаев А.В., Потапов А.С.* Комбинированный метод подсчета эритроцитов на изображениях мазков крови // Научно-технический вестник СПбГУ ИТМО. 2012. № 01(77). С. 19–23.
3. *Miller M.N.* Leukocyte Classification by Morphological Criteria // Eng. Foundation Conf. on Automatic Cytology. 1972. Saxtons River. VT. P. 14–18
4. *Van Osta P., Geusebroek J.M., Ver Donck K., Bols L., Geysen J., ter Haar Romeny B.M.* The principles of scale space applied to structure and colour in light microscopy // Proceedings of the Royal Microscopical Society. 2002. V. 37(3). P. 161–166.
5. *Cheng Ke.* White Blood Cell Detection Using a Novel Fuzzy Morphological Shared-Weight Neural Network // International Symposium on Computer Science and Computational Technology. 2008. V. 2. P. 532–535.
6. *Xubo B. Song, Yaser S. Abu-Mostafa, Joseph Sill, Harvey Kasdan* Incorporating Contextual Information in White Blood Cell Identification // Proceedings of the 1997 conference on Advances in neural information processing systems 10. 1998. MIT Press Cambridge, MA, USA. P. 950–956.
7. *Prof. Samir K. Bandyopadhyay, Sudipta Roy* Detection of Sharp Contour of the element of the WBC and Segmentation of two leading elements like Nucleus and Cytoplasm // International Journal of Engineering Research and Applications. 2012. V. 2. № 1. P. 545–551.
8. *Devesh D. Nawgaje* Implementation of Fuzzy Inference System For White Blood Cell Cancer Detection Using DSP TMS320C6711 // International Journal of Engineering Science and Technology. Special Issue. 2011. P. 123–127.
9. *Giemsa G.* Eine Vereinfachung und Vervollkommnung meiner Methylenazur-Methylenblau-Eosin-Färbemethode zur Erzielung der Romanowsky-Nochtschen Chromatinfärbung // Centralbl f Bakt. 1904. V. 37. P. 308–311.
10. *Otsu N.* A threshold selection method from gray-level histograms // IEEE Trans. Sys., Man., Cyber. 1979. V. 9. P. 62–66.
11. *Samet H., Tamminen M.* Efficient Component Labeling of Images of Arbitrary Dimension Represented by Linear Bintreees // IEEE Transactions on Pattern Analysis and Machine Intelligence. 1988. V. 10. P. 579.
12. *Steinhaus H.* Sur la division des corps materiels en parties // Bull. Acad. Polon. Sci. 1956. C1. P. 801–804.
13. *Lloyd S.* Least square quantization in PCM's // Bell Telephone Laboratories Paper. 1957. P. 129–137.