

СЕГМЕНТАЦИЯ И КЛАССИФИКАЦИЯ ЦВЕТНЫХ ИЗОБРАЖЕНИЙ ГИСТОЛОГИЧЕСКИХ СРЕЗОВ

Мартыненко Т.В.

Донецкий национальный технический университет, г. Донецк
кафедра автоматизированных систем управления

E-mail: mart@kita.dgtu.donetsk.ua

Abstract

Martynenko T.V. "Segmentation and classification of histological cuts color images". Segmentation of bright blobs in an image is an important problem in computer vision and particularly in biomedical imaging. In the article methods of color images segmentation such as region growing, thresholding, k-means and edge detection are considered. Experimental results for negative and positive cells extraction using images containing examples having different types and shapes of cells are presented. The accuracy estimation of classification methods for histological objects is resulted.

Общая постановка проблемы. Многие отрасли медицины и техники, имеющие отношение к получению, обработке, хранению и передаче информации, в настоящее время в значительной степени ориентируются на развитие информационных систем обработки и анализа изображений. Большинство медицинских изображений характеризуются низким качеством, вследствие воздействия различных случайных помех и шумов в процессе их формирования, наличием большого числа сложных многофакторных объектов, что определяет высокие требования к достоверности, надежности и точности результатов обработки. Использование математических методов и проблемно-ориентированных алгоритмов позволяет существенно снизить продолжительность исследования, и вместе с тем повысить достоверность диагностики, основанных на обработке изображений.

Автоматизация анализа морфологии гистологических срезов позволяет ускорить процесс обработки препаратов и повысить достоверность обнаружения изменений внутренних органов при онкологических заболеваниях, а также расширяет возможности профилактики и предупреждения злокачественных новообразований. В настоящее время морфологическая диагностика опухолей (8–10% от общего количества новообразований) нуждается в применении таких методов как: фазово-контрастная микроскопия, гистохимический анализ, электронная микроскопия и т.д., что позволяет оценить гистологический материал по принципу норма-патология.

Одной из основных задач автоматизации обработки гистологических срезов является обнаружение и измерение параметров объектов (размер, текстура, расположение, количество однотипных объектов). Эта задача может быть эффективно решена с помощью методов и средств цифрового анализа изображений.

Сложность получения качественных гистологических препаратов и высокая вариабельность большинства гистологических структур являются основными причинами отсутствия эффективных автоматизированных систем в гистологии. Однако увеличение быстродействия ЭВМ позволяет использовать алгоритмы большой вычислительной сложности, значительно расширить круг исследований, открывает новые пути решения задач, основанных на обработке изображений и анализе сцен.

Постановка задач исследования. Задачей данной работы является разработка алгоритмов распознавания и классификации гистологических объектов на цветных изображениях с целью диагностики злокачественных новообразований.

Для решения задачи требуется реализация следующих подзадач:

- классифицировать изображения гистологических объектов;
- разработать методы сегментации цветных изображений гистологических объектов;
- провести апробацию методов на конкретных примерах.

Объекты гистологических и гистохимических исследований характеризуются широким спектром форм, цветов и размеров, что вносит определенные особенности на этапе сегментации. В работе [1] представлены основные признаки цитологической картины гистологических срезов любой топологии злокачественных опухолей. Окрашивание гистологического препарата позволяет выделить определенные элементы цитологической картины, поэтому для анализа срезов цвет играет важную роль. Следует отметить, что стандартное преобразование цветного изображения в полутоновое приводит к потере некоторых объектов различной окраски, но одинакового уровня яркости.

Алгоритм выделения положительных и отрицательных ядер гистохимического изображения основывается на двух основных процедурах: цветовой сегментации и выделении контуров.

Методы сегментации можно рассматривать как формализацию понятия выделяемости объекта из фона или понятий связанных с градиентом яркости. Алгоритмы сегментации характеризуются некоторыми параметрами надежности и достоверности обработки. Они зависят от того, насколько полно учитываются дополнительные характеристики распределения яркости в областях объектов или фона, количество перепадов яркости, форма объектов и др. В случае автоматизации анализа гистологических срезов сегментацию необходимо рассматривать как основной начальный этап, заключающийся в построении формального описания изображения, качество выполнения которого во многом определяет успех решения задачи распознавания и интерпретации объектов.

В общем случае сегментация представляет собой операцию разбиения конечного множества плоскости, на которой определена функция исходного изображения $f(x,y)$ на k непустых связанных подмножеств s_i ($i=[1,k]$) в соответствии с некоторым предикатом P , определяемом на множестве $S=\{s_1,s_2,\dots,s_k\}$ и принимающий истинные значения, когда любая пара точек из каждого подмножества s_i удовлетворяет некоторому критерию однородности (например, критерий однородности, основанный на оценке максимальной разности яркости отдельного пикселя и среднего значения яркости, вычисленного по соответствующей области) [2].

Методы сегментации можно разделить на следующие основные группы:

1. Пороговые методы сегментации основаны на преобразовании функции яркости изображения оператором вида:

$$\begin{aligned}
 & TH: f(i, j) \rightarrow f_s(i, j), \\
 & f_s(i, j) = \begin{cases} \lambda_p & \text{при } T_p \leq f(i, j) < T_{p+1} \\ \lambda_0 & \text{при } f(i, j) \leq T_0 \\ \lambda_{K-1} & \text{при } f(i, j) > T_{K-1} \end{cases} \quad (1)
 \end{aligned}$$

где $f_s(i,j)$ — сегментированное изображение; K — число областей сегментации; $\lambda_0, \lambda_1, \dots, \lambda_{K-1}$ — метки сегментированных областей; T_0, T_1, \dots, T_{K-1} — упорядоченные величины порогов $T_0 < T_1 < \dots < T_{K-1}$ (при выборе порогов применяют, как правило, гистограмму распределения яркостей цифрового изображения).

2. Методы, основанные на кластеризации. Слияние пикселей в кластеры осуществляется по мере их близости в выбранном пространстве признаков (например, цветовом). Сначала для каждого пикселя определяются координаты в пространстве признаков, далее в пространстве все множество данных разбивается на кластеры.

3. Методы наращивания областей. Методы этой группы основаны на использовании локальных признаков изображения. Идея метода наращивания областей состоит в анализе сначала стартовой точки, затем ее соседних точек и т.д. в соответствии с некоторым критерием однородности, и в последующем зачислении проанализированных точек в ту или иную группу (количество стартовых точек должно быть равно количеству однородных областей на изображении).

4. Методы контурной сегментации изображений можно разделить на следующие основные классы:

- методы высокочастотной фильтрации;
- методы пространственного дифференцирования;
- методы функциональной аппроксимации;

Большинство алгоритмов определения контуров объектов рассматривают границы как область резкого перепада функции яркости изображения $f_s(i,j)$. Основное отличие — используемая математическая модель понятия границы и алгоритм поиска граничных точек.

В соответствии с поставленными задачами к алгоритмам выделения контуров в гистологических изображениях предъявляются следующие требования: выделенные контура должны быть утоньшенными, без разрывов и замкнутыми. Таким образом, процесс выделения контуров усложняется в связи необходимостью применять алгоритмы утоньшения и устранения разрывов.

Каждый из предложенных методов имеет свои преимущества и недостатки. Следует отметить, что вследствие высокой вариабельности и слабой контрастности большинства гистологических структур в настоящий момент отсутствуют автоматизированные системы обработки и анализа гистологических срезов, которые дают достоверные результаты.

На основании предложенного обзора можно заключить, что эффективность указанных методов для сегментации изображений гистологических срезов можно выявить только экспериментальным путем.

Решение задач и результаты исследований.

Применим алгоритм кластеризации к трехмерному пространству цветовых координат изображения гистологического среза с помощью k -средних. Для этого возьмем максимально допустимое количество кластеров $Q=3$: объекты в кластере 1 соответствуют фону, в кластере 2 — клеткам, которые не окрасились при гистохимическом анализе (отрицательные клетки) и в кластере 3 — положительным клеткам.

Для количественного описания цвета пикселя используем систему координат $L^*a^*b^*$, которая обеспечивает более точное представление цветов чем RGB. Пространство $L^*a^*b^*$ включает информацию о значении интенсивности 'L*', значении цветности красно-зеленой и голубо-желтой оси ('a*' и 'b*' соответственно). Отметим, что при анализе цветов используются окрестности сравнительно небольших размеров (3x3 пикселя) и при их вычислении берется усреднение в пространстве 'a*b*'.
 Каждому пикселю (i,j) изображения ставится в соответствие образ (вектор признаков) $V_{ij}=(V_{1ij}, V_{2ij}, V_{3ij})$, где $V_{1ij}, V_{2ij}, V_{3ij}$ — соответствующие значения яркости пикселя в системе координат $L^*a^*b^*$. В пространстве образов выполняется классификация на основе вычисления евклидова расстояния (2) между пикселем и цветовым маркером. Минимальное евклидово расстояние соответствует минимальному расстоянию между рассматриваемыми пикселями и маркером [3].

$$d(v_{ij}, m_l) = \sqrt{\sum_{k=1}^2 (v_{kij} - m_{kl})^2}, \tag{2}$$

где m_l — центр кластера l .

Определяется новое значение центра кластера:

$$m_i = \min \{m_1..m_Q\} \tag{3}$$

Результаты сегментации на основании метода k-средних приведены на рисунке 1.

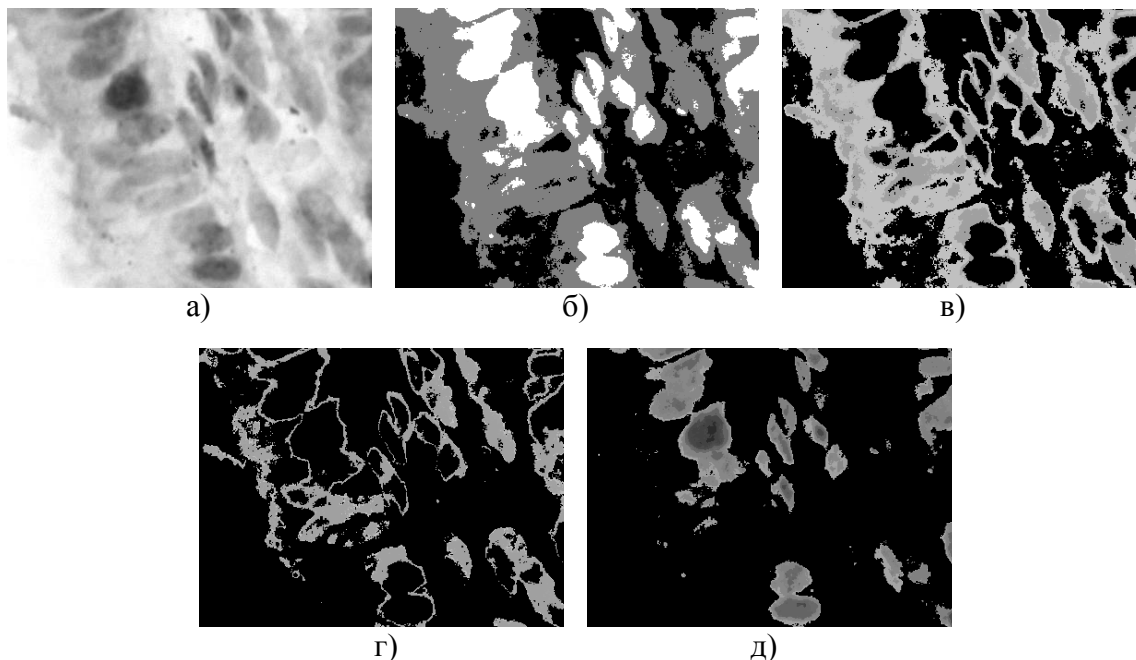


Рисунок 1 — Сегментация гистологического изображения методом k-средних:
 а) исходное изображение; б) отмеченное кластерными индексами;
 в) объекты кластера 1; г) объекты кластера 2; д) объекты кластера 3

Главным достоинством этого подхода является скорость и простота, однако анализ рисунка 1г показывает неудовлетворительное сегментирование отрицательных клеток.

В работе [1] рассматривалось применение различных фильтров для получения контура на изображении. Лучшие результаты показали фильтры Лапласиана-Гауссиана и Канни с различными пороговыми значениями. Для определения границ используется фильтрация исходного изображения I одним из методов. Пиксель считается относящимся к границе, если соответствующий ему пиксель результата фильтрации имеет значение, большее h . Используя функцию edge пакета Matlab, получим результаты обработки изображения выше перечисленными фильтрами по каналам R, G, B (рисунки 2 и 3).

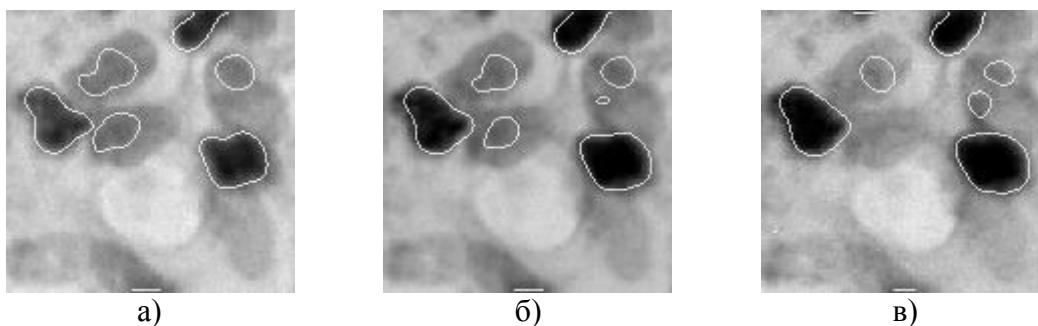


Рисунок 2 — Участки краев, выделенные с помощью фильтра Лапласиана-Гауссиана при $h=5.6$: а) R-канал; б) G-канал; в) B-канал.

Тестирование алгоритмов выделения контуров показало, что в большинстве случаев контуры получаются незамкнутыми и, как следствие, непригодными для ряда процедур анализа. Замкнутый контур дает фильтр Лапласиана-Гауссиана, однако из рисунков 2 и 3 видно, что вне зависимости от выбранного канала большая часть отрицательных клеток не обнаружена.

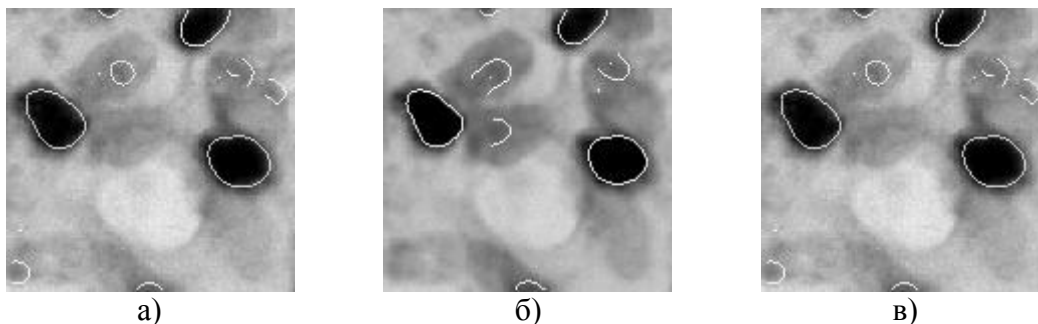


Рисунок 3 — Результаты применения метода оконтуривания Канни при $h=3.9$: а) R-канал; б) G-канал; в) B-канал.

Применим алгоритм наращивания областей к изображению гистологического среза (рисунок 4а). На первом этапе процесса пары квантования элементов изображения объединяются в группы (атомы), если они обладают одинаковой яркостью и являются четырехсвязными [4]. Затем по двум эвристическим правилам удаляются слабые границы между атомами. Две смежные области R_1 и R_2 сливаются, если

$$\frac{D}{\min\{P_1, P_2\}} > \varepsilon_2, \tag{4}$$

где ε_2 — константа, обычно равная 0.5, D — длина той части границы, для которой разность яркостей ΔY граничащих элементов меньше, некоторого коэффициента значимости ε_1 , P_1 и P_2 — периметр смежных областей.

Это эвристическое правило предотвращает слияние смежных областей приблизительно одинакового размера, но допускает поглощение меньших областей большими. По второму правилу происходит ликвидация общих слабых границ, остающихся после применения первого правила. Смежные области сливаются, если

$$\frac{D}{C} > \varepsilon_3, \tag{5}$$

где ε_3 — константа, приблизительно равная 0.75, C — длина общей границы.

Результаты применения метода наращивания областей приведены на рисунке 4.

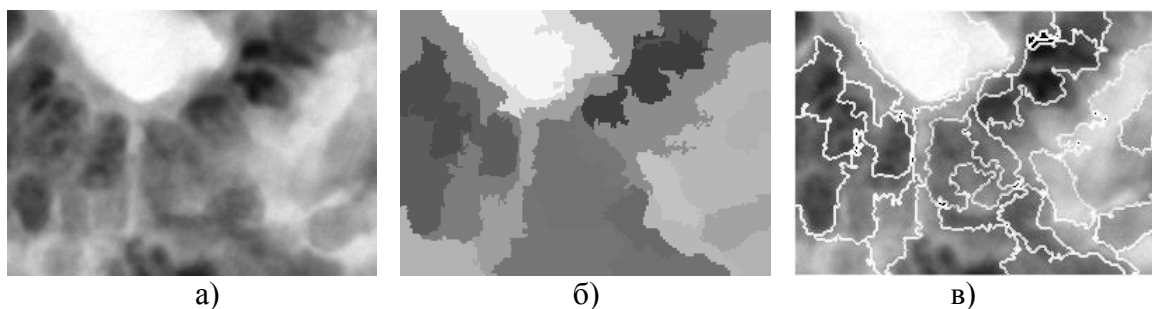


Рисунок 4 — Сегментация изображения гистологического среза методом наращивания областей: а) исходное изображение; б) результат сегментации; в) границы областей

Алгоритмы, основанные на методе наращивания областей, отличаются рядом достоинств: при изменении выбора отображений и меры можно эффективно менять чувствительность алгоритма, высокая точность и скорость нахождения объектов.

Главной проблемой сегментации цветного изображения гистологического среза, основанной на многомерном пороговом ограничении, является выбор пороговой величины. Выбор данной величины основывается на анализе гистограмм яркостей в системах координат цвета RGB и HLS (рисунок 5). Поскольку не все гистограммы признаков являются унимодальными, то определение наилучшего признака и наилучшего уровня, по которому выполняется пороговое ограничение, осуществляется следующим образом [4]:

1. Первым кандидатом является интенсивность, позволяющая разделить изображения на черные и белые области.
2. Следующий кандидат выбирается на основании приоритета, зависящего от уровня контраста и положения мод на гистограмме.

После того как пороговый уровень выбран (для каналов R и B — 160, 220) изображение разбивается на две части. Результаты сегментации цветного гистологического изображения на основании многомерного порогового ограничения представлены на рисунке 6.

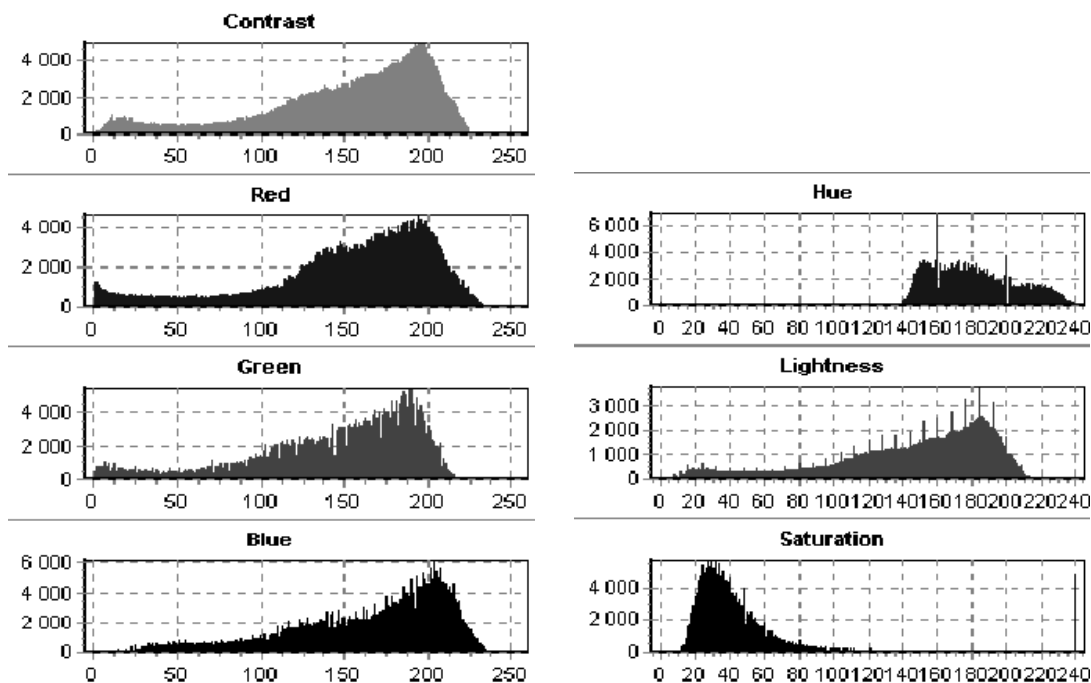


Рисунок 5 — Типичные гистограммы признаков для цветных изображений

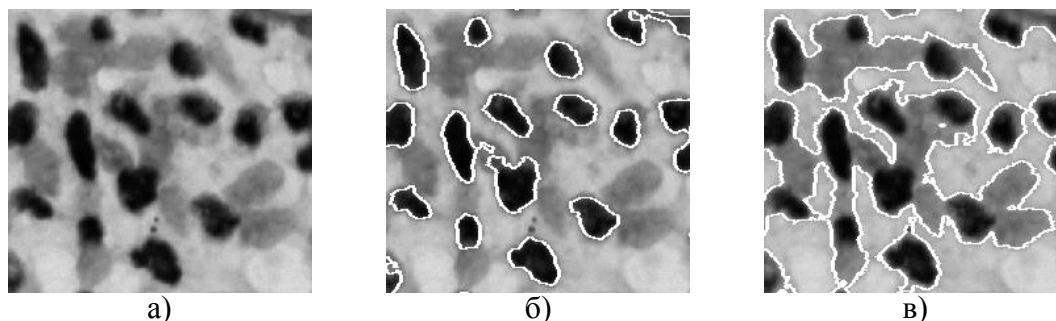


Рисунок 6 — Результат получения контура клеток с помощью цветовой пороговой сегментации: а) исходное изображение; б) положительные клетки; в) отрицательные клетки

Проведенный анализ показывает, что многомерная пороговая сегментация достаточно точно позволяет определить количество положительных клеток и выявить области

отрицательных клеток, требующие дальнейшего разделения (рисунок 6в). Основным недостатком алгоритма является необходимость настроек порогов для гистологических изображений с разными значениями яркости, контрастности или цветового баланса.

Оценка точности автоматической сегментации гистологических объектов указанными методами на реальных изображениях приведена в таблице 1. Эталонные данные для оценки точности работы методов были предоставлены специалистами лаборатории фундаментальных исследований Института Неотложной и Восстановительной Хирургии.

Таблица 1 — Оценка точности автоматической сегментации гистологических объектов

Метод	Время обработки изображения, с	Точность выявления клеток, %	
		положительные	отрицательные
Кластеризация с помощью k-средних	560,31	91,09	67,5
Выделение контуров фильтром Лапласиана-Гауссиана	179,56	83,75	63,94
Наращивание областей	85,93	55,32	40,21
Многомерное пороговое ограничение	4,53	93,75	48,98

Анализ рассмотренных методов показал, что наиболее перспективным, является пороговый метод. Однако, с целью улучшения выделения отрицательных клеток метод требует модификации.

В качестве метода исследования изображений гистологических срезов предлагается использовать модифицированный метод региональных минимумов и максимумов [5], в котором изображение разбивается на вспомогательные сегменты.

Выводы. Выделение гистологических объектов на цифровых изображениях затруднено существенной вариабельностью и слабой контрастностью. Результаты проведенных сравнительных экспериментов по определению эффективности сегментации и классификации объектов гистологических срезов позволяют сделать вывод о возможности использования методов цифровой обработки в анализе изображений, однако рассмотренные методы требуют модернизации. Следовательно, существует необходимость построения взаимосвязанной классификации гистологических объектов и алгоритмов сегментации, а также развития алгоритмов сегментации более общего и универсального характера для выделения широкого класса объектов.

Литература

1. Скобцов Ю.А., Мартыненко Т.В., Белявцев А.А. Обработка изображений гистологических срезов и распознавание опухолевых клеток // Наукові праці Донецького державного технічного університету. Серія: Обчислювальна техніка та автоматизація, випуск 74. — Донецьк: ДонНТУ, 2004р.
2. Яковлев А.В., Сидоренко Е.Н. Методы и аппаратура анализа структуры микрошлифов металлов — Муром. ин-т Владимир. гос. ун-та. — Муром, 2001. — 25 с.; библи. — 18 назв., — Рус., — Деп в ВИНТИ 18.01.2001 #133-B2001 // <http://jakovlev.boom.ru/science.htm>.
3. И.М. Журавель "Краткий курс теории обработки изображений" // <http://matlab.exponenta.ru/imageprocess/book2/index.php>.
4. У.К. Прэтт. Цифровая обработка изображений. — М.: Мир, 1982. Т.1,2. — 792 с.
5. Zakaria Ben Sbeh, Laurent D. Cohen A New Approach of Geodesic Reconstruction for Drusen Segmentation in Eye Fundus Images, IEEE Transactions on medical imaging, vol. 20, no. 12, December 2001.