

МЕТОД ВЫДЕЛЕНИЯ КОНТУРОВ ОБЪЕКТОВ С НЕЧЕТКИМИ КРАЯМИ НА ЦВЕТНЫХ ИЗОБРАЖЕНИЯХ

Медведев Д.Г., начальник компьютерного отдела, ассистент кафедры информатики и прикладной математики Криворожского педагогического института ГВУЗ «Криворожский национальный университет», e-mail: dimakrrog@gmail.com

Ключевые слова: выделение контуров объектов, цветные изображения, размытые границы, сплайн-функции, оптимизация, оператор Соболя.

Проводится краткий обзор методов выделения контуров объектов. Предлагается метод выделения контуров объектов с размытыми границами с использованием аппроксимирующих сплайн-функций.

Введение

Интенсивное развитие средств получения и передачи цифровых изображений создает проблему обработки огромных объемов изображений. Большое количество изображений получают в геологии, минералогии, биологии, металлургии, медицине, метеорологии, экологии и т.п.

Важным источником информации на изображениях является не только наличие или количество определенных объектов, но и их морфология. Например, морфология зерен минералов позволяет определить условия формирования, общности происхождения. Морфология иммунных клеток – эозинофилов – показывает состояние иммунной системы [6]. Однако определение морфологических показателей – трудоемкий процесс, требующий оконтуривания. Эта задача значительно усложняется для микрообъектов, которые по своей природе или из-за особенностей регистрации не имеют четких контуров. Объекты имеют размытые контуры из-за диффузной природы (дым, загрязнения атмосферными выбросами), ввиду «рваного», фрактального характера краев (кроны деревьев, границы больших гранулированных лимфоцитов), из-за разфокусировки изображения.

В оконтуривании наиболее известным и общепринятым является метод активных контуров. Его практический вариант змейка (snake) описан в 1987 году [7]. Функция, моделирующая контур, изменяется, пытаясь достичь минимума собственной энергии (гладкости контура) и несоответствия контура (функция изображения). Метод требует начальной инициализации (как правило 3-5 точек) и в процессе итерационного построения рассчитывает производные и градиенты изображения. Общая идея находит воплощение во многих вариантах метода, отличающихся способами расчета градиентов, производных, методами оптимизации и видом параметрической кривой контура. Разнообразие обусловлено попыткой достичь наибольшей пригодности для конкретных задач, условий и объектов.

Другой подход к оконтуриванию связан с цифровой фильтрацией. Выделение контуров близко к выделению высокочастотной составляющей изображения. Цифровые фильтры выделения контуров реализованы как дискретные дифференциальные операторы 3x3 или 5x5 пикселей. Наиболее известные операторы Роберта, Собеля, Прюитт [4, 5]. Фильтрация изображения этими

операторами позволяет получить контуры изображения во всем поле обработки. Однако по этому изображению трудно определить морфологические показатели отдельных объектов.

Задача оконтуривания значительно усложняется, если контур объекта размыт или недостаточно гладкий. Это приводит к усилению влияния шумов, плохой сходимости итерационного процесса, неточного воспроизведения контуров. Различные авторы, в зависимости от конкретных условий, решают эту задачу по-разному. Например, ведется поиск «поврежденных» фрагментов и настройка на такие фрагменты. Другой метод – метод Hybrid Method использует технологию оконтуривания, при которой выделяются и устраняются по построению контура экстремальные точки. Однако такие алгоритмы требуют дополнительных параметров и их точной настройки, часто при активном участии оператора.

Поэтому задача автоматизации оконтуривания объектов с нечеткими краями за их цифровыми изображениями для определения морфологических параметров и дальнейшей классификации является актуальной.

Анализ последних исследований и публикаций

Одним из перспективных и современных методов выделения контура объекта на изображении является метод активных контуров. Алгоритм заключается в следующем: производится деформация контура с целью заключения внутри его области, которая содержит пиксели объекта [8].

С помощью метода активных контуров можно найти границу объекта на изображениях, в том случае, если начальная граница изображения задается пользователем или другим способом, возможно автоматизированным. Точки в контуре стремятся к границе объекта при минимизации энергии контура. Для каждой отметки в окрестностях v_i , значение энергии находится как:

$$E_i = \alpha E_{int}(v_i) + \beta E_{ext}(v_i) \quad (1)$$

где $E_{int}(v_i)$ является функцией энергии, зависящей от формы контура, и $E_{ext}(v_i)$ – функцией энергии, зависящей от свойств изображения и типа градиента в окрестности точки v_i , α и β являются константами, обеспечивающими относительную коррекцию величин энергии.

E_b , $E_{int}(v_i)$, $E_{ext}(v_i)$ являются матрицами. Значение в центре каждой матрицы соответствует энергии контура в отметке v_i . Другие значения в матрицах соответствуют энергии в каждой отметке в окрестностях v_i . Каждая отметка v_i стремится в v_i' , соответствующей положению минимального значения в E_i . Этот процесс иллюстрируется на рис. 1. Если функции энергии выбраны правильно, то контур V должен приблизиться и остановиться на границе объекта.

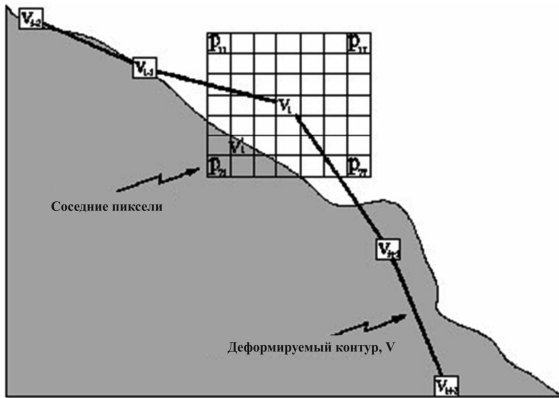


Рис. 1. Пример движения отметки v_i в активном контуре

Недостаток применения метода в автоматизированных системах состоит в определении коэффициентов α и β . Они корректируются оператором. Также оператор выбирает и функции энергий внутри контура и за его пределами.

На основе данного метода разработаны алгоритмы для выделения контуров объектов различной природы. Быстрый алгоритм активных контуров, предложенный Donna J. Williams и Mubarak Shah [4], выделяет контуры различных объектов, однако, как и в самом методе, необходимо прямое участие оператора при выборе коэффициентов.

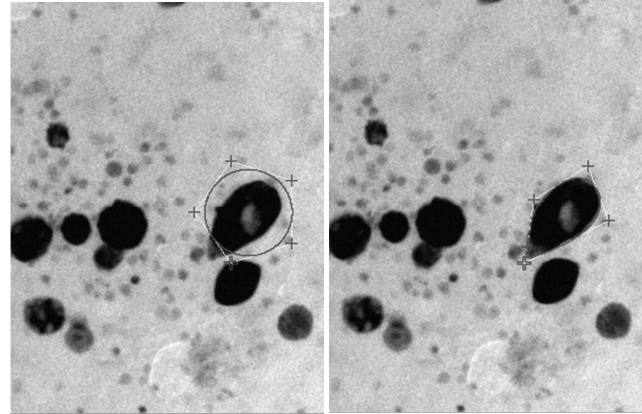
Компанией Biomedical Image Group (BIG) [2] предложен метод моделирования контуров биологических объектов методом змейки (E-Snake) с применением экспоненциальных сплайнов (рис. 2) [3].

Такой алгоритм является наиболее подходящим для морфологического анализа клеток.

Среди различных возможных представлений кривой используется B-сплайн, который характеризуется M точками управления. Точки (крестики) задают контрольный полигон (прямые линии между точками), который управляет сплайн-змейкой (кривая линия).

Эволюция кривой сформулирована как задача минимизации. Связанная с ней функция называется энергией змейки. Используются комбинации контура и области энергий. Энергия контура соответствует интегралу от проекции изображения градиента на кривой. Эта энергия использует сглаженную версию исходного изображения с гауссовским ядром α стандартного отклонения. Области энергий максимизируют разницу между интенсивностью области, ограниченной кривой, и интенсивностью вне ее.

Как отмечают авторы, экспоненциальный сплайн в B-форме лучше всего подходит для моделирования круговых, эллиптических и каплеобразных форм. При аппроксимации минимизируют разность энергий области, ограниченной кривой, и выделенной области. Для минимизации этого показателя используется итерационная процедура оптимизации контрольных точек, определяющих полигон, описанный вокруг объекта моделирования.



а) до оптимизации б) после оптимизации

Рис. 2 Выделенный контур клетки экспоненциальным сплайном

Однако процедура оптимизации несколько осложнена тем, что контрольные точки движутся на плоскости. Точки не имеют прямой связи со строящимся контуром.

При использовании частотных фильтров для выделения контуров чаще всего пользуются операторами Робертса, Собеля, Прюитт. Операторы Собеля и Прюитт используют градиент, вычисленный по формуле $G = \sqrt{G_x^2 + G_y^2}$. Отличие состоит лишь в способе вычисления производных по составляющим. Оператор Робертса использует окно 2×2 , и величина градиента вычисляется по формуле:

$$G_R(j, k) = \sqrt{[F(j, k) - F(j + 1, k + 1)]^2 + [F(j, k + 1) - F(j + 1, k)]^2} \quad (2)$$

Вместе с тем, при использовании данных методов выделяются контуры всех объектов на изображении. Количество найденных контуров зависит от выбранного порогового значения. Если его выбрать высоким, то получается мало контуров, и искомый объект может быть невыделен. При низком значении – большое количество контуров. То есть, в общем случае необходимо участие оператора для выбора порогового значения, что усложняет процесс автоматизации выделения контура.

Автором были выполнены попытки применить данные методы к оконтуриванию объектов с размытыми контурами, но результаты оказались неудовлетворительными (рис. 3).

Цель исследований – разработать метод выделения контура объектов с нечеткими, размытыми краями на цветных изображениях, лишенный в той или иной степени отмеченных выше недостатков.

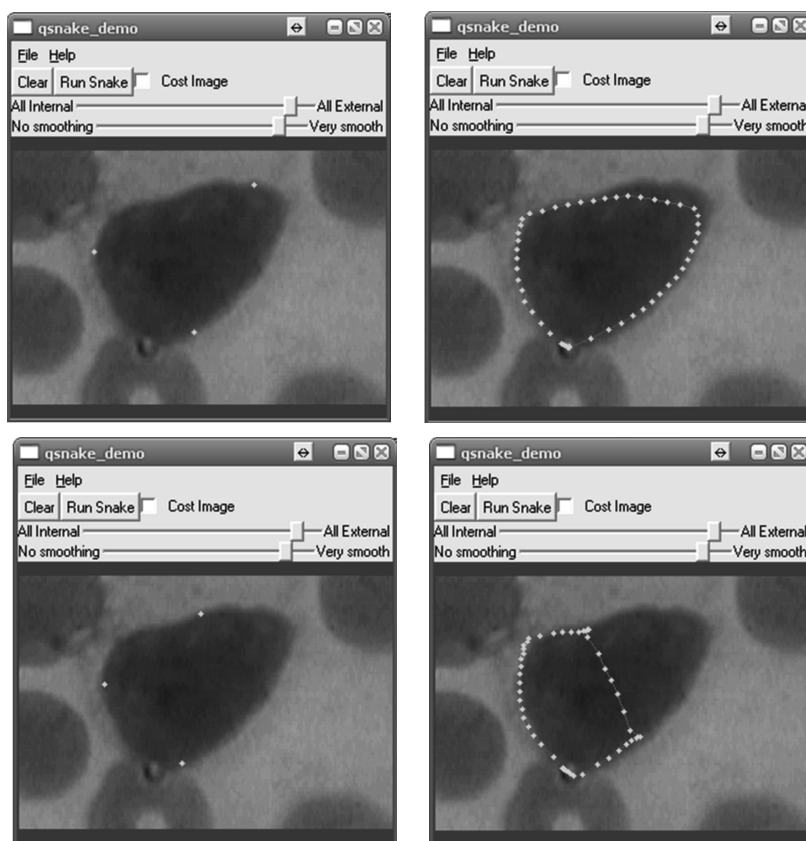


Рис. 3. Применение Qsnake к оконтуриванию эозинофила

Результаты исследований и разработки

В работе предлагается информационная технология оконтуривания объектов с использованием метода, состоящего из последовательных этапов.

Целью предварительной обработки является улучшение качества получаемого изображения. Снимки, как правило, недостаточно четкие и контрастные, а также несут в себе много лишней информации, прежде всего шумов (пыль, осевшую на стекла мазков, другие факторы). Улучшить четкость и контрастность можно аппаратными методами, например, заменой видеокамеры на ту, которая специализируется на съемке изображений под микроскопом. Но это дорогостоящий способ, который не решает проблему исключения шумов. Поэтому целесообразнее воспользоваться алгоритмическими методами улучшения качества изображения.

Уменьшение изолированных препятствий с минимальным размыванием контура обеспечивается за счет одномерной медианной фильтрации по пяти точкам. Перепады яркости на изображении содержат в себе информацию о контурах объектов. Именно с их помощью выполняется разделение объектов друг от друга. Перепады яркости обычно разделяют друг от друга две области изображения с различной яркостью. Идеальный детектор перепада должен указывать наличие перепадов в единой точке, расположенной строго в центре градиента яркости. Для грубого выделения контура применен оператор Собеля с окном 3*3, элементы которого показаны на рис 4.

Расчеты выполняются по формуле:

$$G(j,k) = \sqrt{X^2 + Y^2}, \quad (3)$$

где $X = (A_2 + 2A_3 + A_4) - (A_0 + 2A_7 + A_6)$,

$Y = (A_0 + 2A_1 + A_2) - (A_6 + 2A_5 + A_4)$

A_0	A_1	A_2
A_7	$F(j,k)$	A_3
A_6	A_5	A_4

Рис. 4. Окно расчетов оператором Собеля

Однако, в общем случае, это не позволяет достичь оконтуривания именно необходимого объекта. Кроме того, оператор Собеля требует выбора порогового значения. Для работы оператора изображение переводится в формат YUV. Оператор применяется к сигналу яркости Y.

После использования этих методов получаем границы объектов – бинарное изображение, но границы находятся неоднозначно, как множество отдельных точек. Последнее объясняется нечеткими и размытыми границами тела.

Учитывая только яркость, невозможно выделить нужный контур и определить его важные характеристики. Поэтому необходимо принимать во внимание и компоненты цветности U (сине-желтая) и V (красно-голубая). Для этого к бинарному изображению яркости (компонента Y) добавим компоненты цветности в неизменном виде. Получаем цветное изображение с другим распределением цветов за счет бинарного преобразования яркости. На нем четко выделены те же контуры, но объект стал белого цвета, а фон имеет желтоватый цвет.

Благодаря разнице в цвете выделяем объект. Для этого разработана стратегия последовательного перебора. Находятся приблизительные границы «белого тела». Находятся белые точки, вокруг которых есть доста-

точное количество (не менее 5) таких же белых точек. Это исключает отдельные белые точки среди цветного фона. Далее двигаемся в четырех направлениях до черных точек. В результате получаем декартовы координаты точек контура искомого объекта и прямоугольник, в который объект вписан. Это очень грубое выделение контура, по которому сложно определить морфологические параметры.

Для точного оконтуривания объекта, представленного множеством отдельных точек, предложен новый метод итерационного приближения контура сплайном. Особенность искомого контура в том, что он является замкнутой функцией. Построить циклический, замкнутый сплайн можно двумя способами: переходом к параметрической функции или полярной системе координат.

При использовании параметрической функции появляется сложность выбора параметра t , и имеем две отдельные функции $x(t)$ и $y(t)$. Поэтому был избран перевод в полярную систему координат.

Исследуемые объекты (микроорганизмы, минералы) имеют форму, близкую к выпуклой, что упрощает переход, как будет показано далее на рис. 8-10.

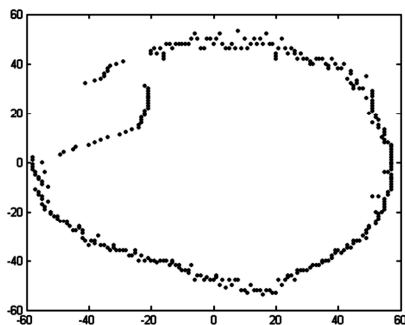
Началом отсчета полярной системы координат служит центр прямоугольника, в который вписан найденный контур (рис. 5).

Для уточнения грубо выделенного контура необходимо построить сплайн-модель. Особенность данного сплайна в том, что он является замкнутым. В таком сплайне количество узлов равно количеству фрагментов. Особенностью периодического сплайна является то, что перед первым фрагментом находится последний, а за последним фрагментом следует первый. Для построения сплайн-модели был выбран сплайн Эрмита. Для вычислений базисных функций сплайна Эрмита используются следующие формулы [2]:

$$\begin{aligned} X_0(x) &= -p_1 p_4 b, & X_1(x) &= p_1 c - p_2 d + b, \\ X_2(x) &= -p_4 c - p_3 d + a, & X_3(x) &= -p_3 p_2 d, \\ a &= (x - x_j) / h_j, & b &= 1 - a, & c &= ab^2, & d &= a^2 b, \\ p_1 &= \frac{h_j}{h_{j-1}}, & p_2 &= \frac{h_j}{h_{j+1} + h_j}, & p_3 &= \frac{h_j}{h_{j+1}}, & p_4 &= \frac{h_j}{h_{j-1} + h_j}, \\ h_{j-1} &= x_j - x_{j-1}, & h_j &= x_{j+1} - x_j, & h_{j+1} &= x_{j+2} - x_{j+1}. \end{aligned} \quad (4)$$

Замкнутость контура сплайна учитывается следующим образом:

- для первого: $h_{j-1} = x_R - x_{R-1}$;
- для последнего: $h_j = x_j - x_0, h_{j+1} = x_2 - x_1$;
- для предпоследнего: $h_{j+1} = x_l - x_0$,



где R – количество фрагментов сплайна, j – номер текущего фрагмента.

В общем виде, для сплайна, который состоит из R фрагментов, имеем выражение:

$$S(x) = \sum_{j=0}^R f(x_j) X_j(x) \quad (5)$$

где $X_j(x)$ локальная базисная сплайн-функция, состоящая из четырех фрагментов:

$$X_j(x) = \begin{cases} X_{0,j-1}(x), & x \in [x_{j-1}, x_j); \\ X_{1,j}(x), & x \in [x_j, x_{j+1}); \\ X_{2,j+1}(x), & x \in [x_{j+1}, x_{j+2}); \\ X_{3,j+2}(x), & x \in [x_{j+2}, x_{j+3}); \\ 0, & x \notin [x_{j-1}, x_{j+3}). \end{cases} \quad (6)$$

Значение сплайна в произвольной точке, принадлежащей j -му фрагменту, равно:

$$S(x) = f(x_{j-1})X_{0,j-1}(x) + f(x_j)X_{1,j}(x) + f(x_{j+1})X_{2,j+1}(x) + f(x_{j+2})X_{3,j+2}(x) \quad (7)$$

Для замкнутого сплайна это выражение истинно для всех фрагментов.

Аппроксимация выполняется по методу наименьших квадратов (МНК). Необходимо на множестве точек $y_i, a_i, i = \overline{1, N}$, оценить истинное значение контура $f(a_i)$ при условии аддитивной погрешности с нулевым математическим ожиданием $y_i = f_i + \varepsilon_i, M[\varepsilon_i] = 0$. Погрешность считаем некоррелированной и распределенной по нормальному закону с дисперсией $D[\varepsilon] = \sigma^2$. Необходимо найти такие значения вектора $A = \{a_{ij}\}$, при которых достигается минимум $d = \sum_{i=1}^N (y_i - S(a_i, A))^2 \rightarrow \min$.

В общем виде МНК с моделью сплайном не отличается от аналогичных уравнений для алгебраических полиномов.

Решением является:

$$A = (P^T P)^{-1} (P^T Y) \quad (8)$$

где P – матрица планирования, столбцы – базисы сплайна (6).

Особенностью сплайн-модели контура является его замкнутость. Начало и конец сплайна являются условными, так как конечный узел одновременно является и первым. Поэтому матрица планирования для замкнутого сплайна имеет блочно-диагональную форму с дополнительными блоками справа-вверху и слева-внизу (рис. 6).

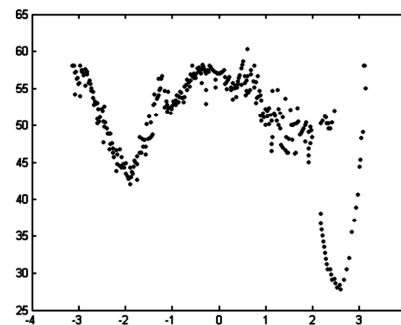


Рис. 5. Функция контура в декартовой и полярной системе координат

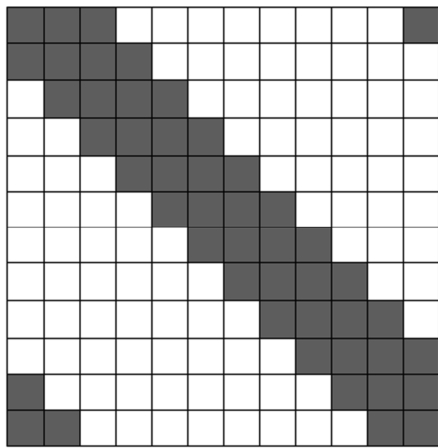


Рис. 6 Схема расположения блоков матрицы планирования

Для построения матрицы планирования P используются формулы (4). Для каждой i -ой точки вычисляются значения функций $X_0(x_i)$, $X_1(x_i)$, $X_2(x_i)$, $X_3(x_i)$. В матрице планирования они находятся в i -ой строчке. Для каждого j -го фрагмента с общего количества фрагментов R , кроме $j = 1, R - 1, R$, значения функций $X_0(x_i)$, $X_1(x_i)$, $X_2(x_i)$, $X_3(x_i)$ находятся в столбцах с номерами $j - 1, j, j + 1, j + 2$ соответственно. Для первого фрагмента ($j = 1$) значения функций $X_0(x_i)$ записываются в столбец R , так как в качестве нулевого фрагмента используется последний. Значения $X_1(x_i)$, $X_2(x_i)$, $X_3(x_i)$ записываются в столбцы 1, 2, 3 соответственно. Для предпоследнего фрагмента ($j = R - 1$) значения $X_3(x_i)$ заносятся в 1-ый столбец, так как за последним фрагментом идет первый. Значения $X_0(x_i)$, $X_1(x_i)$, $X_2(x_i)$ записываются в столбцы $R - 2, R - 1, R$ соответственно. По той же причине для последнего фрагмента ($j = R$) значения $X_3(x_i)$ заносятся во 2-й столбец, $X_2(x_i)$ заносятся в 1-ый столбец. Значения $X_0(x_i)$, $X_1(x_i)$, записываются в столбцы $R - 1, R$ соответственно.

Гладкость и точность приближения сплайн-контура неявно зависит от количества узлов сплайна. Для достижения приближения сплайном в $\sigma < 3$ разработан метод оптимизации количества узлов и их размещения. На начальной стадии сплайн строится на четырех фрагментах с четырьмя узлами. Первый узел находится в первой точке сплайна. Остальные три – на равных расстояниях друг от друга. Уточнение положения узлов сплайна выполняется следующим образом. Расстояние между соседними узлами делится на пять частей. Каждый узел движется в каждую из них. Если среднеквадратическое отклонение сплайн контура после временного движения узла уменьшилась, то отныне узел находится в этой новой точке. Если после оптимизации размещения узлов не удается достичь необходимой точности приближения, то увеличиваем число фрагментов сплайна (добавляем еще один узел): в середину фрагмента с наибольшим отклонением добавляется новый узел. Далее вновь повторяется оптимизация размещения узлов. В среднем сплайн строится на 10-12 узлах. Результат оптимизации показан на рис. 7.

Разработанный метод можно применять при оконтуривании объектов различной природы. С его помощью были выделены контуры на цифровых изображениях минераграфии (рис. 8), металлографии (рис. 9), иммунологии (рис. 10).

Предложенный метод реализован в информационной технологии определения морфологических параметров эозинофилов. Исследуя морфологические признаки больших гранулированных лимфоцитов (эозинофилов), можно определить иммунный статус человека, то есть стадию борьбы организма с вирусами. Морфологические признаки эозинофилов здорового и больного человека существенно отличаются. Для определения иммунного статуса пациента применен классический дискриминантный анализ [8].

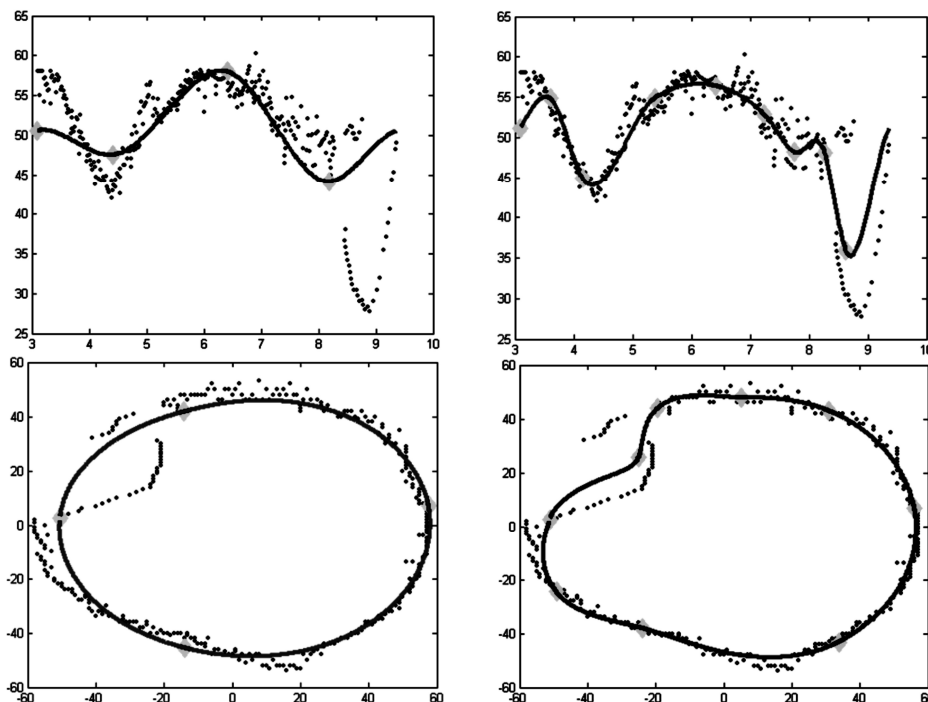


Рис. 7. Контур объекта до и после оптимизации

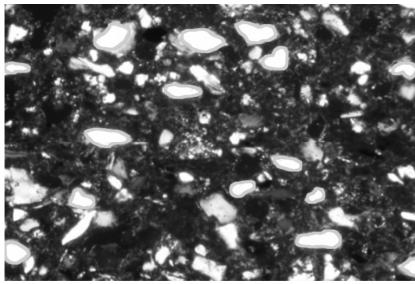


Рис. 8. Контуры срезов минералов

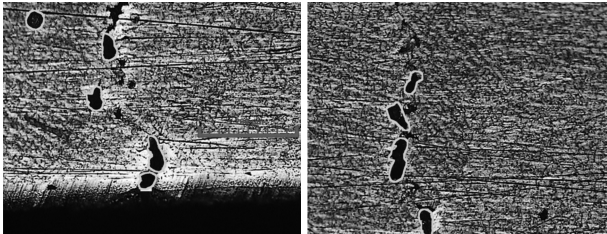


Рис. 9. Контуры поврежденной металлической поверхности

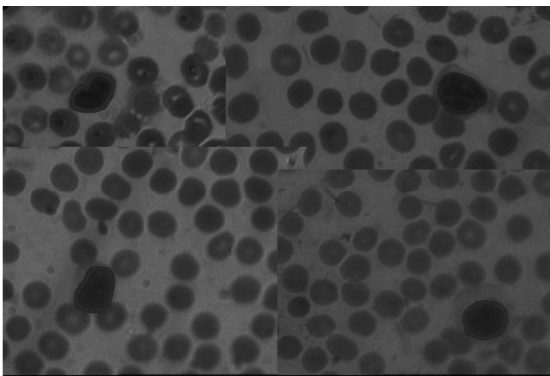


Рис. 10. Контуры объектов микроскопии

Для определения дискриминантной функции использованы относительные показатели: $x_1 = \frac{d_v}{d_m}$,

$$x_2 = \frac{S}{l^2 / 4\pi}, \quad x_3 = \frac{l}{2\sqrt{\pi S}},$$

где d_v – большой диаметр, d_m – малый диаметр, S – площадь, l – длина контура. Для расчета использована обучающая выборка из $n = 110$ образцов, из них $n_1 = 72$, $n_2 = 38$.

Определены коэффициенты β_i дискриминантной функции и получена дискриминантная функция $Y(X) = \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3$.

Вероятность ошибочной классификации определяется следующим образом. Если m_1 – число наблюдений W_1 , отнесенных к W_2 , и m_2 – число наблюдений W_2 , классифицированных в W_1 , то $P(2|1) = m_1/n_1$ и $P(1|2) = m_2/n_2$.

По данным ретроспективного анализа было получено $P(2|1) = 0,073$, а $P(1|2) = 0,126$. Это вероятность отношения здорового пациента к больным, а больного – к здоровым.

Заключение

Разработанный метод, работа которого проверялась на реальных изображениях, продемонстрировал устойчивость и достижение заданной точности.

Метод включает в себя последовательность следующих операций.

1. Преобразование изображения в YUV.
2. Медианная фильтрация.
3. Определение грубого контура оператором Собеля.
4. Определение координат контура.
5. Аппроксимация сплайном.
6. Оптимизация контура.

Небольшое нарушение выпуклости объектов не является препятствием для предложенного метода, хотя в отдельных случаях может несколько снижать точность за счет сглаживания.

Литература

1. Шелевицкий І.В., Шутко М.О., Шутко В.М., Калганова О.О., Сплайни в цифровій обробці даних і сигналів. Кривий Ріг: Видавничий дім. 2008. – 232с., іл.
2. Biomedical Imaging Group: Gradient (index) <http://bigwww.epfl.ch/algorithms.html>
3. DONNA J. WILLIAMS AND MUBARAK SHAH A Fast Algorithm for Active Contours and Curvature Estimation. CVGIP: IMAGE UNDERSTANDING Vol. 55, No. 1, January, pp. 14-26, 1992
4. Roberts L.G. Machine Perception Of Three-Dimensional Solids / L.G. Roberts – Massachusetts Institute of Technology Lincoln Laboratory. Режим доступа: <http://www.packet.cc/files/mach-per-3D-solids.html>
5. Samuel J. Dwyer III. A personalized view of the history of PACS in the USA. / Samuel J. Dwyer III. // Proceedings of the SPIE, «Medical Imaging 2000: PACS Design and Evaluation: Engineering and Clinical Issues». Edited by G. James Blaine and Eliot L. Siegel, 2000. – P. 2-9.
6. Литвинова Т.В. Профілактика респіраторних захворювань у дітей, хворих на бронхіальну астму, що одержують базисну терапію: дис. канд. мед. наук: 14.01.10 / Т.В. Литвинова. – Д.:ДДМА, – 2006. – 193с.
7. Kaas M., Witkin A., Terzopoulos D. Snakes: Active Contour Models. // Int. Journal of Computer Vision. – 1987, N1, p.312-331.
8. Факторный, дискриминантный и кластерный анализ / Дж.-О. Ким, Ч.У. Мьюллер, У.Р. Клекка и др.; под ред. И.С. Енюкова. Пер. с англ.– М.:Финансы и статистика, 1989. – С. 90-93.

THE METHOD OF EDGE DETECTION OF OBJECTS WITH FUZZY EDGES IN COLOR IMAGES

Medvedev D. G.

A brief review of methods of edge detection of objects is considered. Basic methods do not allow to allocate the contours of objects with blurred edges or require the active participation of the user. The advantage of the proposed method of edge detection using the approximating spline functions is its automatization (the user only need to specify one point belonging to the object). The result is a contour in the form of the analytic function that simplifies parameterization. The developed method can be used in edge detection of objects of different nature: mineragraphy, metallo-graphy, immunology. The method is implemented in the information technology for determination of morphological parameters of eosinophils.