

УДК 681.513

ПРИМЕНЕНИЕ ПРИНЦИПОВ ПОСТРОЕНИЯ СИСТЕМ МАШИННОГО ЗРЕНИЯ В ЗАДАЧЕ АНАЛИЗА ИЗОБРАЖЕНИЙ КЛЕТОЧНЫХ СТРУКТУР

Ковригин А.В., – аспирант

Политехнический институт Сибирского федерального университета

Рассматривается подход на основе принципов построения систем машинного зрения для решения задачи анализа изображений клеточных структур. Преимущество предлагаемого подхода состоит в возможности обойти проблему повышения ошибки классификации при наличии пересечений объектов с полупрозрачной структурой.

Ключевые слова: МАШИННОЕ ЗРЕНИЕ, КОМПЬЮТЕРНОЕ ЗРЕНИЕ, АВТОМАТИЧЕСКАЯ КЛАССИФИКАЦИЯ, СЕГМЕНТАЦИЯ, КЛАСТЕРИЗАЦИЯ

1. Введение

Большинство медицинских методик ручного микроскопического анализа эксплуатирует свойства глаз человека распознавать соотношения формы, оттенков, текстуры сложных изменчивых объектов без их измерения, на основе качественных характеристик. Если исследуемых клеток много, а требуемый объем выборки невелик (100-200), то визуальный анализ (например, подсчет лейкоформулы) при наличии достаточного числа опытных врачей-лаборантов может выполняться в массовом порядке. Значительно хуже справляются глаза человека с дифференциальным подсчетом более серьезных выборок клеток (около 500), сопровождающимся большим количеством измерений (например, определение индекса овалоцитоза эритроцитов). Весьма трудны психологически и физически для человека задачи поиска редких клеток, подсчета сложных дифференциальных формул, аккуратного просмотра больших пространств препарата (например, подсчет формул при цитопении, поиск юных клеток в мазке крови, подсчет миелограммы в мазке костного мозга, подсчет патологических типов эритроцитов). Очень трудны и даже вредны для глаз врача анализы с подсчетом

флуоресцирующих меток (например, подсчет субпопуляции меченых FITC лимфоцитов). Далеко не вся присутствующая в цитологическом препарате диагностическая информация может быть извлечена глазами врача.

В связи с этим доминировавшая до последнего времени ручная микроскопия даже при высокой квалификации врача-лаборанта тяжела, субъективна и ограничена по своим возможностям.

Развитие алгоритмов машинного зрения позволяет постепенно перейти от ручного способа анализа клеточных структур. Принципы построения системы машинного зрения подразумевают наличие трех уровней (фильтрация, сегментация, классификация), каждому из которых соответствует свой набор алгоритмов. Однако если существующий набор алгоритмов на нижнем уровне (фильтрация) позволяет качественно выполнить все требования, выдвигаемые к данному уровню при решении поставленной задачи, то решения, существующие на среднем уровне (сегментация), недостаточны. Так традиционные алгоритмы сегментации изображений, основанные на принципе близости точек объекта по цвету (яркости) и не учитывающие пересечения объектов, исключают возможность последующей классификации с приемлемой для рассматриваемой задачи ошибкой.

В работах Ковригина А.В., Хлопова А.Н. [1, 2] рассматривался синтез алгоритмов автоматической классификации объектов на изображении. На основе синтезированных алгоритмов и принципов построения системы машинного зрения, в настоящей статье предложен подход к анализу изображений, содержащих полупрозрачные структуры клеток, позволяющий снизить ошибку классификации, в том числе и при наличии геометрического пересечения.

2. Постановка задачи исследования

Задана исходная выборка $v = (x^i, y^i, a^i, i = \overline{1, n})$, представляющая собой цифровой снимок клеточных структур, где x^i, y^i, a^i – соответственно абсцисса, ордината и яркость i -ой точки изображения.

Требуется определить количество разнотипных клеток на снимках.

3. Анализ изображений, содержащих полупрозрачные структуры клеток

Идея предлагаемого подхода основывается на построении системы классификации изображения на принципе построения системы машинного зрения.

На первом уровне выполняется предобработка исходной выборки, с целью повышения качества распознавания. Предобработка включает в себя пересчет значений яркостей a^i на основе выравнивания гистограммы яркостей исходной выборки (1).

На втором уровне используется принцип алгоритма автоматической классификации цветных изображений [1], основанный на построении R окрестности вокруг точек (x^i, y^i, a^i) выбранного класса Ω и проверки гипотезы (2) о принадлежности точек (x^i, y^i, a^i) , попавших в R окрестность, классу Ω_m .

Третий уровень включает в себя логический контроль на однородность класса Ω_m (4) и, в случае неоднородности, построение карты минимальных расстояний по точкам класса Ω_m (5) с последующим их перераспределением между дополнительно выявленными классами [2].

Предлагаемая методика:

Уровень 1:

1. Вычислить число точек выборки соответствующих каждому из уровней яркости

$$h_l = \sum_{i=1}^n h(a^i, l),$$

где $l = \overline{0,255}$ - уровень яркости,

$$h(a^i, l) = \begin{cases} 1, a^i = l \\ 0, a^i \neq l \end{cases} - \text{единичная функция, определяющая равенство}$$

значения яркости i -ой точки выборки V и уровня яркости l .

2. Вычислить «выровненные» значения яркости

$$\bar{a}^i = \frac{255}{256 * 256} \sum_{j=1}^{a^i} h_j \quad (1)$$

3. Сформировать из выборки V и значений \bar{a}^i выборку

$$\bar{V} = (x^i, y^i, \bar{a}^i, i = \overline{1, n}).$$

Уровень 2:

1. Выбрать i -ую точку из выборки \bar{V} и отнести её к классу Ω_m . Включить точку (x^i, y^i, \bar{a}^i) в подвыборку V^m m -го класса, с последующим её удалением из выборки \bar{V} .
2. Определить R окрестность класса Ω_m из условия попадания в нее соседних точек (x^j, y^j, \bar{a}^j) .
3. Проверить все точки (x^j, y^j, \bar{a}^j) из R окрестности на принадлежность их классу Ω_m , путем применения решающего правила

$$m_1(x) = \begin{cases} (x^i, y^i, \bar{a}^i) \in \Omega_m \quad \forall \|\bar{a}^i - \bar{a}^j\| \leq z \\ (x^i, y^i, \bar{a}^i) \notin \Omega_m \quad \forall \|\bar{a}^i - \bar{a}^j\| > z \end{cases}, \quad (2)$$

где \bar{a}^i - значение яркости точки (x^i, y^i, \bar{a}^i) , принадлежащей классу Ω_m ,

\bar{a}^j - значение яркости точки (x^j, y^j, \bar{a}^j) , принадлежащей R окрестности,

$\|\bar{a}^i - \bar{a}^j\| = \sqrt{3|\bar{a}^i - \bar{a}^j|} \leq z$ - мера близости между точками (x^i, y^i, \bar{a}^i) и (x^j, y^j, \bar{a}^j) ,

$z = 5$ - коэффициент размытости изображения.

При условии принадлежности точки (x^j, y^j, \bar{a}^j) из R окрестности классу Ω_m включить (x^j, y^j, \bar{a}^j) в выборку V^m m -го класса, с последующим ее удалением из \bar{V} .

4. При отсутствии точек из R окрестности, которые можно отнести к классу Ω_m , следует увеличить счетчик классов на единицу $m = m + 1$

При отсутствии точек из выборки \bar{V} , которые можно отнести к классу Ω_m при заданном коэффициент размытости изображении z , следует перейти к этапу 5, в противном же случае перейти к этапу 1.

5. Установить индекс текущего класса $k = 1$
 6. Построить R окрестность вокруг точек класса Ω_k и проверить гипотезу о принадлежности каждой точки (x^j, y^j, \bar{a}^j) из этой окрестности классу Ω_k , путем применения решающего правила

$$m_2(x) = \begin{cases} (x^j, y^j, \bar{a}^j) \in \Omega_k \forall \|\bar{a}^j - M_{\Omega_k}(a)\| < \|\bar{a}^j - M_{\Omega_r}(a)\| \\ (x^j, y^j, \bar{a}^j) \in \Omega_r \forall \|\bar{a}^j - M_{\Omega_k}(a)\| \geq \|\bar{a}^j - M_{\Omega_r}(a)\| \end{cases} \quad (3)$$

где $M_{\Omega_k}(a) = \frac{1}{|\Omega_k|} \sum_{\Omega_k} \bar{a}_k$ - среднее значение яркости по классу Ω_k

$$M_{\Omega_r}(a) = \frac{1}{|\Omega_r|} \sum_{\Omega_r} \bar{a}_r - \text{среднее значение яркости по классу } \Omega_r$$

r - индекс класса, к которому принадлежит точка (x^j, y^j, \bar{a}^j)

При условии перехода точки из одного класса в другой по правилу (3), необходимо пересчитать средние показатели этих классов. Для этого следует применять рекуррентные формулы.

$$M_{\Omega_k} = \frac{n-1}{n(n-1)} \sum_{i=1}^{n-1} \bar{a}^i + \frac{1}{n} \bar{a}^n = M_{\Omega_k} + \frac{1}{n} \bar{a}^n,$$

$$M_{\Omega_r} = \frac{n-1}{n(n-1)} \sum_{i=1}^{n-1} \bar{a}^i - \frac{1}{n} \bar{a}^n = M_{\Omega_r} - \frac{1}{n} \bar{a}^n,$$

где Ω_k - класс, к которому перешла точка (x^j, y^j, \bar{a}^j) , Ω_r - класс, который она покинула.

7. При выполнении условия $k + 1 \leq m$, увеличить индекс текущего класса на 1 и перейти к этапу 6, в противном же случае перейти к этапу 8.
 8. Увеличить коэффициент размытости изображения z на величину dz и перейти к этапу 5.

Величину dz выбирать равной 3, пока z не достигнет 20, далее рекомендуется dz выбирать равным 1.

9. Если значение z достаточно велико ($z \approx [30, 35]$) следует перейти к выполнению уровня 3.

Уровень 3:

1. Для каждого класса Ω_k , $k = \overline{1, m}$, вынести решение об однородности путем применения решающего правила (4)

$$m_3(x) = \begin{cases} s_i^2 > I * \overline{s^2}, & \text{класс } \Omega_k \text{ однороден} \\ \text{иначе,} & \text{класс } \Omega_k \text{ неоднороден} \end{cases}, \quad (4)$$

где s_i^2 - значение дисперсии точек, принадлежащих классу Ω_m ,
 $\overline{s^2}$ - среднее по всем классам Ω_k , где $k = \overline{1, m}$, значение дисперсии,

$I \approx 0.8$ - заранее выбранное пороговое значение.

При условии, что все классы однородны, следует прекратить выполнение алгоритма.

2. Для каждой i -ой точки выборки V_k , где k - индекс неоднородного класса Ω_k , вычислить евклидово расстояние до всех точек, не принадлежащих выборке V_k , и выбрать минимальное

$$d^i = \min \sqrt{(x^i - x^j)^2 + (y^i - y^j)^2}, \quad (5)$$

где x^i, y^i - координаты точки $(x^i, y^i, a^i) \in V_k$,

x^j, y^j - координаты точки $(x^j, y^j, a^j) \notin V_k$.

3. Сформировать из выборки V_k и значений d^i выборку $V_k' = (x^i, y^i, \overline{a^i}, d^i, i = \overline{1, n_1})$.
4. Установить начальное пороговое значение $T = \max d^i$ и счетчик дополнительно выявленных классов $p = 1$.
5. Выбрать i -ую точку из выборки V_k' , имеющую значение $d^i \geq T$, отнести её к классу Ω_{m+p} .
6. Определить R окрестность как набор точек $(x^j, y^j, \overline{a^j}, d^j)$ из выборки V_k' , имеющих значение $d^j \geq T$ и непрерывно связанных с классом Ω_{m+p} .

7. Отнести все точки $(x^j, y^j, \bar{a}^j, d^j)$ из R окрестности к классу Ω_{m+p} и увеличить счетчик дополнительно выявленных классов на единицу $p = p + 1$.
8. При наличии точек из выборки V_k' , которые можно отнести к классу Ω_{m+p} следует вернуться к пункту 12.
9. Уменьшить пороговое значение на единицу $T = T - 1$.
10. Для каждого из классов Ω_l , где $l = \overline{m, m+p}$, определить R окрестность как набор точек $(x^j, y^j, \bar{a}^j, d^j)$ из выборки V_k' , имеющих значение $d^j \geq T$ и граничащих с классом Ω_l . В случае если точка $(x^j, y^j, \bar{a}^j, d^j)$ из выборки V_k' граничит одновременно с несколькими классами, ее следует удалить из выборки V_k' .
11. Отнести все точки $(x^j, y^j, \bar{a}^j, d^j)$ из R окрестностей к соответствующим классам с последующим удалением этих точек из выборки V_k' .
12. При наличии точек из выборки V_k' , которые можно отнести к одному из классов Ω_l , где $l = \overline{m, m+p}$, следует вернуться к пункту 17.
13. При наличии точек в выборке V_k' с значением $d^j \geq T$, следует вернуться к пункту 12.
14. При условии $T > 0$ и наличии не расклассифицированных точек из выборки V_k' следует вернуться к пункту 16.
15. Удалить класс Ω_k из списка найденных классов и уменьшить счетчик классов на единицу $m = m - 1$.
16. При отсутствии неоднородных классов Ω_k , где $k = \overline{1, m}$, следует прекратить выполнение алгоритма, в противном случае следует вернуться к пункту 2.

4. Экспериментальные данные

С целью проверки эффективности предлагаемого подхода, в качестве тестовых примеров, были использованы цифровые снимки мазка периферической крови человека.

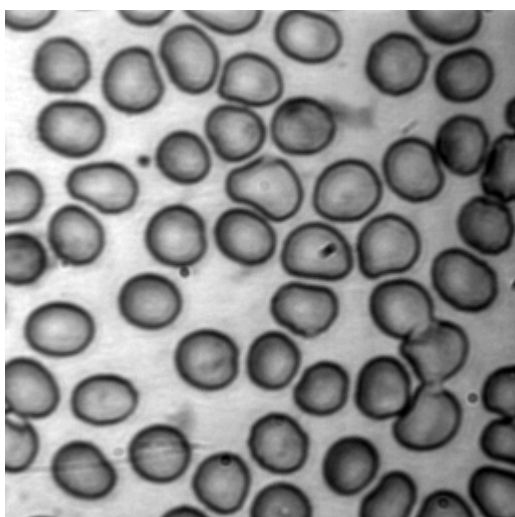
На сухое обезжиренное предметное стекло наносили 4 мкл периферической крови. Мазок производили шлифованным стеклом под углом 30-45°, высушивали на воздухе. Фиксировали мазки в метаноле в течении 5 мин., высушивали на воздухе.

Окраску производили по методу Романовского-Гимза. Высушенные мазки помещали в кювету с рабочим раствором краски (1:40) на 45 мин. Промывали проточной водой и высушивали на воздухе.

Далее разные участки полученного мазка фотографировали камерой Sony SSC-M370CE через тринокулярный микроскоп МБС-9 с увеличением $\times 64$ (40 \times 1,6), полученные изображения размером 255 \times 255, 8 bit(grey) обрабатывались созданными программными средствами.

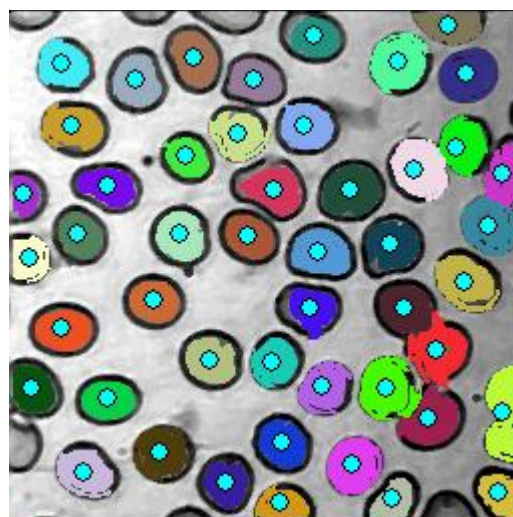
Анализ на основе тестовой выборки, включающей в себя 30 изображений, показал, что оценка вероятности ошибки составила 0.07.

а



в

б



г

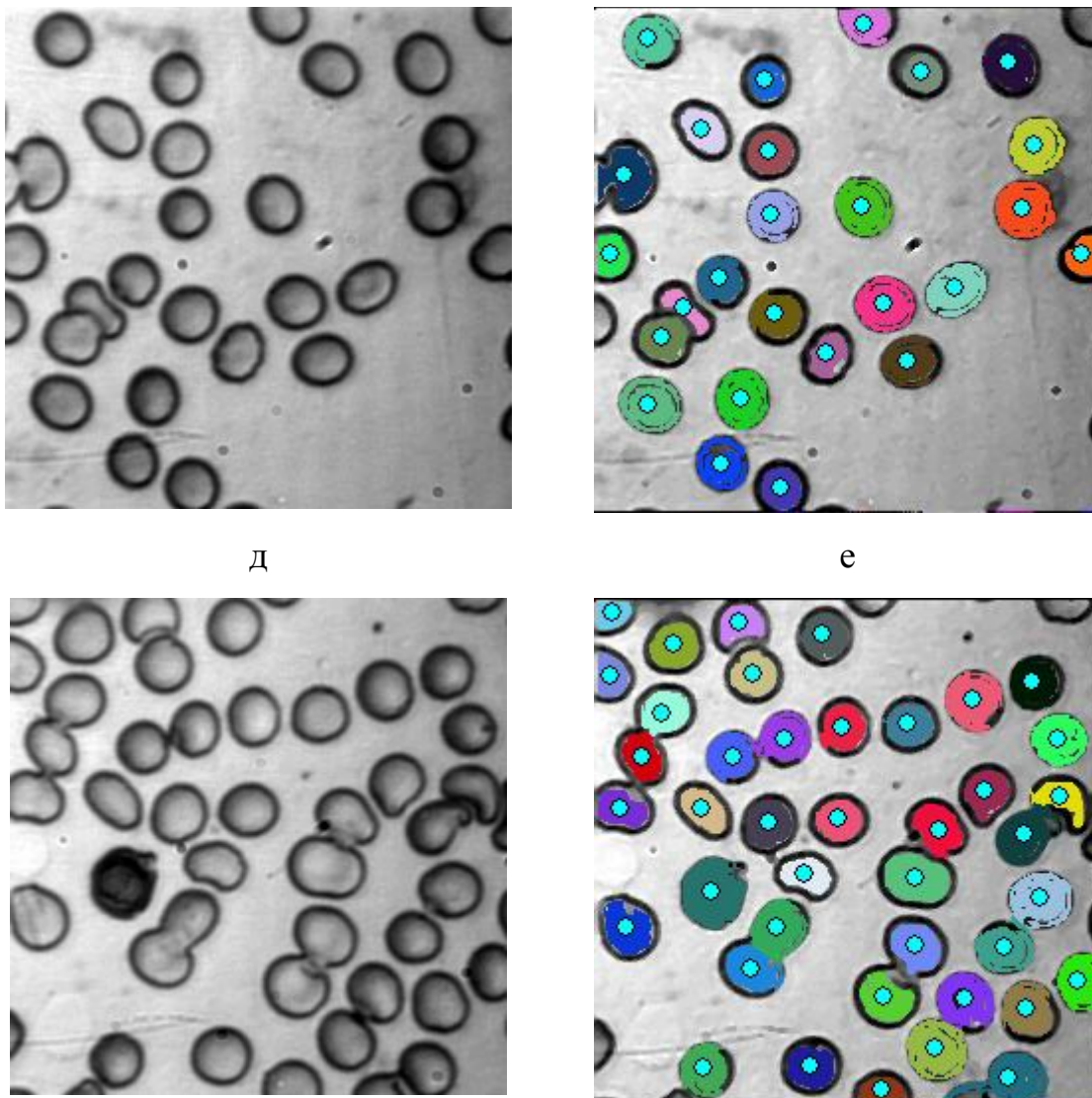


Рисунок 1 – Примеры работы предложенного подхода (а, в, д – исходные фотографии мазков крови; б, г, е – результат работы предложенного подхода).

5. Заключение

Анализ предложенного подхода на основе принципов построения системы машинного зрения показал преимущества перед использованием традиционных алгоритмов автоматической классификации изображений, позволяя автоматизировать работу по исследованию клеточных структур и отказаться от метода ручного микроскопического анализа. Важное достоинство данного подхода заключено в снижении степени зависимости

качества решения задачи от цвета, прозрачности, формы и расположения в пространстве объектов.

Литература

1. Ковригин А.В., Хлопов А.Н. Синтез и анализ алгоритмов автоматической классификации цветных изображений // Электронный журнал "Исследовано в России". 2005. с. 2417-2422. <http://zhurnal.ape.relarn.ru/articles/2005/236.pdf>
2. Ковригин А.В. Многоуровневый алгоритм автоматической классификации изображений в условиях отсутствия четких границ между объектами // Вестник КрасГАУ. Красноярск, 2006, выпуск № 12, с. 407-411
3. Форсайт Д., Понс Ж. Компьютерное зрение. Современный подход. : Пер. с англ.- М.: Издательский дом «Вильямс», 2004. -928 с.: ил. – Парал.тит.англ.
4. Непараметрические системы классификации / А.В. Лапко, В.А. Лапко, М.И. Соколов, С.В. Ченцов. – Новосибирск: Наука, 2000. – 240 с.
5. Лапко А.В. Непараметрические методы классификации и их применение. - Новосибирск: ВО «Наука», 1993.-152 с
6. Автоматизированная подсистема распознавания и оконтуривания клеток / А.А. Коков, Е.В. Меркулова, В.Г. Адамов, В.Я. Спорыхин // Наукові праці ДонДТУ, Донецьк, 2001, випуск 25, с. 110-116
7. Спорыхин В.Я. Автоматизация процесса исследования функционального состояния клеток / В.Я. Спорыхин, Е.В. Адамова // Наукові праці ДонДТУ, Донецьк, 2001, випуск 25, с. 110-116
8. Спорыхин В.Я. Автоматизированная система определения жизнеспособности клеток / В.Я. Спорыхин, Е.В. Адамова. // Наукові праці ДонНТУ, Донецьк, 2002, випуск 38, с. 110-114