

РАЗРАБОТКА МЕТОДОВ ФОРМИРОВАНИЯ ЦВЕТО-ТЕКСТУРНЫХ ПРИЗНАКОВ ДЛЯ АНАЛИЗА БИМЕДИЦИНСКИХ ИЗОБРАЖЕНИЙ

А.И. Пластинин^{1,2}, А.В. Куприянов^{1,2}, Н.Ю. Ильсова^{1,2}

¹Институт систем обработки изображений РАН, Самара, Россия,

²Самарский государственный аэрокосмический университет, Самара, Россия

Аннотация

В статье рассматривается метод формирования цвето-текстурных признаков. Метод основан на вычислении характеристик многомерного распределения вероятности интенсивностей цветовых компонент. Представлены экспериментальные исследования метода на натуральных изображениях. Приводится сравнение разработанного метода с уже существующими.

Введение

Одним из важнейших диагностических методов, который до сих пор выполняется вручную лаборантами, остается общеклиническое исследование крови. Именно данный анализ крови тонко отражает реакцию кровяных органов на воздействие на организм различных физиологических и патологических факторов [1]. Важным диагностическим показателем является количество и процентное соотношение каждого из типов лейкоцитов (лейкоцитарная формула). Лейкоциты различны по внешнему виду и функциям. Различают следующие виды лейкоцитов: нейтрофилы (палочко-ядерные, сегментно-ядерные), эозинофилы, базофилы, моноциты, лимфоциты [2].

Широкое распространение, трудоемкость выполнения и низкое качество результатов при ручном анализе делает автоматизацию общеклинического анализа крови необходимой и актуальной задачей. Рассмотрим принцип работы системы автоматической диагностики. Изображение препарата крови формируется подключенной к окуляру микроскопа фотокамерой. Оно представляет собой фон, с относительно постоянной яркостью, на котором присутствуют различные форменные элементы (клетки крови). Затем на изображении выделяются лейкоциты, которые представляют собой связанные области. Для каждого лейкоцита определяется его тип. И подсчитывается общее количество клеток каждого типа.

Целью работы является разработка признаков для классификации изображений лейкоцитов. Рассматриваются классы: эозинофилы, базофилы, моноциты, лимфоциты.

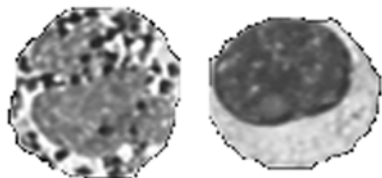


Рис. 1. Изображения лейкоцитов

В качестве основы для исследований был выбран метод, предложенный Хараликом в [3], называемый в дальнейшем статистическим текстурным анализом. Задачей является выделение признаков, связанных с локальными текстурными особенностями

лейкоцитов. Подход, основанный на матрицах вероятностного распределения (МВР) яркости, уже не раз очень успешно использовался для выделения аналогичной информации [4, 5, 6].

К недостаткам методов МВР можно отнести: работу с полутоновыми изображениями; использование целочисленного формата данных; высокую вычислительную сложность, заключающуюся в необходимости работы с трёхмерными матрицами распределения вероятности. В данной работе рассматривается модификация существующих методов для устранения указанных недостатков.

1. Метод цвето-текстурного анализа

Текстурные признаки – это статистические характеристики многомерного вероятностного распределения интенсивности отсчетов цветовых компонент.

Будем рассматривать цветное изображение в формате RGB, которое представляется отсчетами функции интенсивности цветовых компонент $f(m, n) = (f_R(m, n) f_G(m, n) f_B(m, n))^T$, где $m, n \in \mathbb{Z}$. Как известно в пространстве RGB присутствует кросс-канальная корреляция, что приводит к статистической зависимости признаков, вычисленных по разным цветовым каналам. Поэтому необходимо использовать декоррелированное цветовое пространство. Таким является цветовое пространство Ruderman $l\alpha\beta$ [7], которое получено на основе анализа главных компонент в пространстве LMS [7].

Переход в цветовое пространство $l\alpha\beta$ из RGB осуществляется в два этапа.

1) Переход из RGB в LMS [7]:

$$\begin{pmatrix} L \\ M \\ S \end{pmatrix} = \begin{pmatrix} 0,3811 & 0,5783 & 0,0402 \\ 0,1967 & 0,7244 & 0,0782 \\ 0,0241 & 0,1288 & 0,8444 \end{pmatrix} \begin{pmatrix} R \\ G \\ B \end{pmatrix};$$

2) Переход из LMS в $l\alpha\beta$ [7]:

$$\begin{pmatrix} l \\ \alpha \\ \beta \end{pmatrix} = \begin{pmatrix} 0,5774 & 0,5774 & 0,5774 \\ 0,4082 & 0,4082 & -0,8164 \\ 0,7071 & -0,7071 & 0 \end{pmatrix} \begin{pmatrix} \lg L \\ \lg M \\ \lg S \end{pmatrix}.$$

Здесь цветовые компоненты RGB нормированы так, чтобы принимать значения в диапазоне $[0;1]$.

Для того, чтобы достичь инвариантность к цвету, необходимо привести статистические характеристики каждой компоненты к заранее определенным величинам, например, $m = 0$; $\sigma = 1$. Это можно сделать, выполнив для каждого канала поэлементную обработку следующего вида:

$$l' = \frac{l - m^l}{\sigma^l}; \quad \alpha' = \frac{\alpha - m^\alpha}{\sigma^\alpha}; \quad \beta' = \frac{\beta - m^\beta}{\sigma^\beta},$$

где m – математическое ожидание, σ – дисперсия.

В данной работе рассматриваются статистические методы формирования признаков, основанные на 3-х мерной функции распределения вероятности яркости в 3-х соседних точках.

Рассмотрим интенсивности заданного цветового канала C (одного из $\{l, \alpha, \beta\}$) в трех точках, находящихся на заданном расстоянии друг от друга:

$$M_0 = f_c(m, n), \quad M_1 = f_c(m + \Delta m_1, n + \Delta n_1),$$

$$M_2 = f_c(m + \Delta m_2, n + \Delta n_2).$$

Обозначим $P(M_0 = x, M_1 = y, M_2 = z)$ – вероятность появления точек с яркостями x, y, z . Тогда функция распределения вероятности будет определена как: $F^1(x, y, z) = P(M_0 < x, M_1 < y, M_2 < z)$.

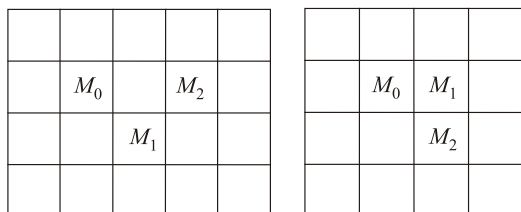


Рис. 2. Примеры взаимного расположения точек

Так как нет необходимости учитывать порядок расположения соседних отсчетов, перейдем к следующей функции распределения, усреднением по следующей схеме:

$$F^2(x, y, z) = \frac{1}{6} (F^1(x, y, z) + F^1(x, z, y) + F^1(y, x, z) + F^1(y, z, x) + F^1(z, x, y) + F^1(z, y, x)).$$

Для обеспечения инвариантности, относительно поворота на угол кратный 45° . Будем рассматривать следующие расположения точек:

1. $\Delta m_1 = \Delta k_1; \Delta n_1 = \Delta l_1; \Delta m_2 = \Delta k_2; \Delta n_2 = \Delta l_2,$
2. $\Delta m_1 = \Delta l_1; \Delta n_1 = \Delta k_1; \Delta m_2 = \Delta l_2; \Delta n_2 = \Delta k_2,$
3. $\Delta m_1 = \Delta l_1; \Delta n_1 = \Delta k_1; \Delta m_2 = -\Delta l_2; \Delta n_2 = \Delta k_2,$
4. $\Delta m_1 = -\Delta k_1; \Delta n_1 = \Delta l_1; \Delta m_2 = -\Delta k_2; \Delta n_2 = \Delta l_2.$

Пусть каждому из случаев соответствуют функции распределения $F_1^2(x, y, z), F_2^2(x, y, z), F_3^2(x, y, z), F_4^2(x, y, z)$.

В результате усреднения получим функцию распределения:

$$F(x, y, z) = \frac{1}{4} (F_1^2(x, y, z) + F_2^2(x, y, z) + F_3^2(x, y, z) + F_4^2(x, y, z)).$$

Введем следующие числовые характеристики:

1. Корреляция:

$$R = \int_{R^3} xyz dF(x, y, z).$$

2. Затенение:

$$S = \int_{R^3} (x + y + z) dF(x, y, z).$$

3. Контраст:

$$K_1 = \frac{1}{3} \int_{R^3} (|x - y| + |x - z| + |y - z|) dF(x, y, z),$$

$$K_2 = \int_{R^3} |(x - y)(x - z)(y - z)| dF(x, y, z).$$

4. Инерция:

$$I_1 = \frac{1}{3} \int_{R^3} ((x - y)^2 + (x - z)^2 + (y - z)^2) dF(x, y, z),$$

$$I_2 = \int_{R^3} (x - y)^2 (x - z)^2 (y - z)^2 dF(x, y, z),$$

Признаки записаны для нулевого математического ожидания. Общий вид признака, можно представить как:

$$C = \int_{R^3} g(x, y, z) dF(x, y, z).$$

В случае, если исходная функция распределения не известна, можно использовать эмпирическую функцию распределения вероятности [8]:

$F(x, y, z) = \frac{k}{N}$, где k – количество элементов выборки, попадающих в область $x' < x, y' < y, z' < z$. Тогда получим следующую выборочную оценку:

$$\tilde{C} = \frac{1}{N} \sum_{i=0}^{N-1} g(x_i, y_i, z_i).$$

Значение признаков вычисляется для каждой цветовой компоненты, поэтому далее будем использовать обозначения, включающие название цветового канала и признака, например: l Shading (lS) – затенение, рассчитанное по компоненте l.

Таким образом, процедура вычисления признаков заключается в одном проходе по изображению заданными шаблонами взаимного расположения точек. Экспериментальные исследования на натуральных изображениях показали, что предложенный метод вычисления признаков является более быстрым и требует меньше памяти по сравнению с методом, основанным на вычислении 3-х мерной матрицы вероятностного распределения (гистограммы).

2. Дискриминантный анализ признаков

Для определения степени эффективности признаков при классификации используются элементы дискриминантного анализа [9, 10]. Цель выбора признаков состоит в выделении признаков, которые являются наиболее эффективными с точки зрения разделимости классов. Разделимость классов зависит только от распределения объектов в классах, а также от используемого классификатора.

В дискриминантном анализе критерии разделимости записывают с использованием матриц рассеяния внутри классов и матриц рассеяния между классами. Пусть S_w , – матрица рассеяния внутри классов, S_b – матрица рассеяния между классами, S_m – матрица рассеяния смеси. Для выбора признаков рассматривались три критерия:

$$J_1 = \text{tr } S_m / \text{tr } S_w,$$

$$J_2 = |S_m| / |S_w|,$$

$$J_3 = \text{tr}(S_w^{-1} S_b).$$

Наилучшие признаки те, для которых критерий разделимости максимален [9, 10]. В данной работе рассматривалась задача выбора трёх признаков. Из всего множества векторов признаков, построенных по обучающей выборке, вычислялись критерии разделимости по всем тройкам, количество которых равно $C_{18}^3 = 816$. В таблицах 1-3 показаны результаты вычисления критериев разделимости.

Таблица 1. Значения критерия 1

lS	αK_2	βS	10,66213
αS	αK_2	βS	10,66211
...
αR	αS	βS	2,08106

Таблица 2. Значения критерия 2

lS	αK_2	βR	163,95
lS	αK_2	βS	159,86
...
lK_2	lI_2	αR	3,38

Таблица 3. Значения критерия 3

lS	αK_2	βS	17,09
lI_1	αS	αK_2	17,01
...
lK_2	lI_2	αR	1,80

Из таблиц видно, что критерии J_1 и J_3 дали одинаковые результаты, при этом наилучшими оказались признаки: $l\text{ Shading}$ (lS), $a\text{ Contrast2}$ (αK_2), $b\text{ Shading}$ (βS) – 1-ый набор. А по критерию J_2 лучшими признаками являются: $l\text{ Shading}$ (lS), $a\text{ Contrast2}$ (αK_2), $b\text{ Correlation}$ (βR) – 2-ой набор.

Расположения объектов в пространстве признаков данных наборов показано на рисунках 2, 3, из которых видно, что обрабатываемые четыре класса лейкоцитов линейно разделимы.

Можно заметить, что полученные наборы признаков отличаются одним признаком (признак βS и βR). Вследствие этого, в дальнейшем для построения классификатора планируется использовать четыре признака, максимизирующие выбранные критерии разделимости классов.

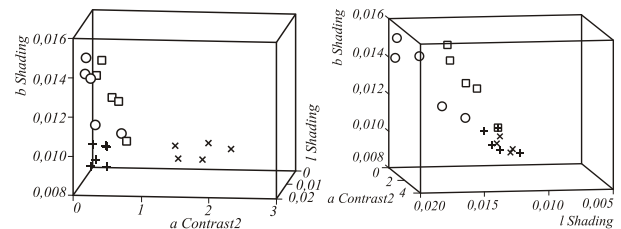


Рис. 3. Расположение объектов в пространстве 1

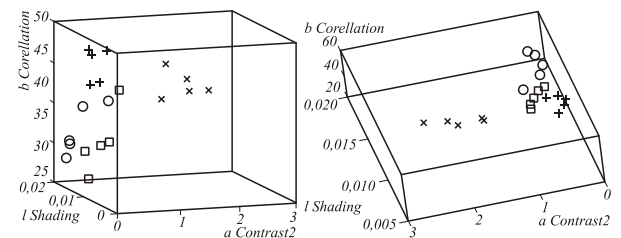


Рис. 4. Расположение объектов в пространстве 2

Заключение

В работе предложен метод вычисления цветотекстурных признаков, основанный на вычислении статистических характеристик двумерных случайных полей. Экспериментальные исследования проводились на натуральных изображениях клеток крови. Работа алгоритма сравнивалась с уже реализованными алгоритмами на основе МВР. В результате сделаны следующие выводы:

1. Полученные признаки позволяют проводить автоматическую классификацию клеток крови.
2. Алгоритм обладает высокой скоростью, в сравнении с методами, основанными на анализе матрицы вероятности [5, 6].

Анализ проведенных экспериментальных исследований свидетельствует о том, что разработанные алгоритмы могут применяться для системы диагностики заболеваний крови. Одним из достоинств работы можно назвать сокращение затрат времени на обработку изображения.

Благодарности

Работа выполнена при поддержке российско-американской программы «Фундаментальные исследования и высшее образование» (BRHE) и программы Президиума РАН «Фундаментальные науки – медицине», гранта РФФИ № 06-07-08006-офи, гранта РФФИ № 07-08-96611.

Литература

1. Kern W. PDQ Hematology. – B.C. Decker. Published October, 2002. – 440 p.
2. Thieml H. Color Atlas of Hematology, 2nd revised ed. – Thieme, 2004. – 198 p.
3. R.M.Haralick, K.Shanmugam, I.Dinstein Textural features for image classification // IEEE Trans.Syst.Man Cybern. 1973. – V. 3. – P. 610-621.
4. Shirvaikar M., Trivedi M. Developing texture-based image clutter measures for object detection // Optical Engineering, 1992. – Vol. 31(12). – P. 2628-2639.
5. Basset O., Dupont F., Hernandez A., Odet C. Texture image analysis: application to the classification of bovine muscles from meat slice images // Optical. Engineering, 1999. – V. 38(11). – P. 1950-1959.
6. Hozman J., Bernas M., Kubinek R. Biomedical Image Processing Applications // Biomedical Engineering & Biotechnology, Publishing House of the Czech Technical University, 1994.
7. Ruderman D.L., Cronin T.W. and Chiao C.C. Statistics of Cone Responses to Natural Images: Implications for Visual Coding // J. Optical Soc. of America, 1998. – Vol. 15. – No. 8. – Pp. 2036-2045.
8. Айвазян С.А., Енюков И.С., Мешалкин Л.Д. Прикладная статистика: основы моделирования и первичная обработка данных /– М.: Финансы и статистика, 1983. – 471 с.
9. Фукунага К. Введение в статистическую теорию распознавания образов. – М.: Наука, 1979. – 270 с.
10. Theodoridis S., Koutroumbas K. Pattern Recognition, 2nd ed // Academic press, 2003.

DEVELOPMENT OF METHODS OF COLOR AND TEXTURE CHARACTERS FORMING FOR BIOMEDICAL IMAGES ANALYZING

A.L. Plastinin^{1,2}, A.V. Kupriyanov^{1,2}, N.Yu. Ilyasova^{1,2}

¹*Image Processing Systems Institute of the RAS, Samara, Russia,*

²*Samara State Aerospace University (SSAU), Samara, Russia*

Abstract:

This paper presents a method of color and texture characters forming. The method is based on calculating characteristics of multivariate distribution of rate probability for color components. Experimental studies of the method are given for full-scale images. The developed method is compared with those ones already exist.

Keywords: biomedical images, color and texture features, multivariate distribution of rate probability

Citation: Plastinin AL, Kupriyanov AV, Ilyasova NYu. Development of methods of color and texture characters forming for biomedical images analyzing [In Russian]. Computer Optics 2007; 31(2): 82-85.

Acknowledgements: The work was partly supported by the Russian-American Basic Research and Higher Education Program (BRHE), the “Basic Sciences for Medicine” Basic-Research Program of the Presidium of the Russian Academy of Sciences and by the Russian Foundation for Basic Research (RFBR) grants Nos. 06-07-08006-ofi and 07-08-96611.

References:

- [1] Kern W, Decker BC. PDQ Hematology, 2002; 440 p.
- [2] Thieml H. Color Atlas of Hematology, 2nd revised ed. Thieme, 2004; 198 p.
- [3] Haralick RM, Shanmugam K, Dinstein I. Textural features for image classification. IEEE Trans.Syst.Man Cybern. 1973; 3: 610-621.
- [4] Shirvaikar M, Trivedi M. Developing texture-based image clutter measures for object detection. Optical Engineering, 1992; 31(12): 2628-2639.
- [5] Basset O, Dupont F, Hernandez A, Odet C. Texture image analysis: application to the classification of bovine muscles from meat slice images. Optical. Engineering 1999; 38(11): 1950-1959.
- [6] Hozman J, Bernas M, Kubinek R. Biomedical Image Processing Applications. Biomedical Engineering & Biotechnology. Publishing House of the Czech Technical University, 1994.
- [7] Ruderman DL, Cronin TW, Chiao CC. Statistics of Cone Responses to Natural Images: Implications for Visual Coding. J. Optical Soc. of A 1998; 15(8): 2036-2045.
- [8] Ayvazyan SA, Enyukov IS, Meshalkin LD. Applied statistics: basics of modeling and preprocessing of data [In Russian]. Moscow: Finances and Statistics Publisher, 1983; 471 p.
- [9] Fukunaga K. Introduction to statistical pattern recognition [In Russian]. Moscow: “Nauka” Publisher, 1979; 270 p.
- [10] Theodoridis S, Koutroumbas K. Pattern Recognition, 2nd ed. Academic press, 2003.