

# Воспроизведение люминесцентных изображений раковых клеток с использованием имитационных моделей

Шитик М.М., Лисица Е.В., Яцков Н.Н., Апанасович В.В.  
Кафедра системного анализа и компьютерного моделирования  
Белорусский государственный университет  
Минск, Республика Беларусь  
e-mail: maxshitik@mail.ru

**Аннотация**—В работе предложен алгоритм для имитационного моделирования синтетических изображений раковых опухолей. Разработанная модель может быть рекомендована для исследования и отладки алгоритмов автоматических методов сегментации в цитометрии.

**Ключевые слова:** раковые клетки; люминесцентные изображения; имитационная модель.

## I. ВВЕДЕНИЕ

Качественный и количественный анализ данных в цитометрии до применения автоматических алгоритмов был весьма трудоемким и субъективным процессом [1]. Вполне закономерной необходимостью является появление автоматического анализа, во многом упрощающего процесс обработки данных. В сложившейся ситуации моделирование синтетических изображений может значительно повысить эффективность анализа, путем улучшения методов обработки и их исследовании на смоделированных наборах данных. Варьирование параметров моделирования позволяет исследовать алгоритмы анализа на устойчивость к различным факторам, оказывающим наиболее существенное влияние на проведение реального эксперимента. Конечно, моделирование изображения со свойствами, максимально приближенными к реальным, является очень сложным и все еще нерешенным вопросом, однако пренебрежение некоторыми вторичными деталями объектов позволяет сконцентрировать внимание на основных свойствах объектов и измерительной системы. Более того, упрощения просто необходимы при моделировании таких сложных систем, как клетки, многообразие видов и характеристик которых требует большой гибкости от моделирующей системы.

## II. ИМИТАЦИОННАЯ МОДЕЛЬ

В представленной модели процесс получения изображения разделяется на последовательные стадии, соответствующие реальному проведению эксперимента при помощи люминесцентного микроскопа. На первом этапе генерируется идеальное изображение, состоящее из маркированных специальным красителем клеток [1]. Далее полученное изображение искажается с учетом погрешностей, вносимых измерительной системой: неравномерное освещение объекта, автофлуоресценция фона, оптические погрешности, шум от фотозлектронного умножителя (ФЭУ) и многое другое. Таким образом, на выходе получается изображение с теми же

свойствами, что и у реальных люминесцентных изображений [2].

## III. МОДЕЛИРОВАНИЕ ЕДИНИЧНОЙ КЛЕТКИ

Основным компонентом моделирования изображения является одиночная клетка. В данном методе для генерации формы используется параметрическая модель, в которой форма задается как полигон с заданным количеством вершин около некоторой окружности. Дальнейшее преобразование полигона связано с перемещением вершин полигона и соответствующим изменением формы объекта. Положение  $i$ -ой вершины определяется при помощи следующих уравнений:

$$\begin{aligned}x_i(\theta_i) &= r[U(-\alpha, \alpha) + \cos(\theta_i + U(-\beta, \beta))] \\y_i(\theta_i) &= r[U(-\alpha, \alpha) + \sin(\theta_i + U(-\beta, \beta))],\end{aligned}\quad (1)$$

где  $U(a,b)$  – равномерное распределение на интервале  $[a,b]$ ,  $\theta$  – полярный угол, параметр  $\beta$  отвечает за случайность разбиения, а  $\alpha$  – за случайность радиуса.

Для таких сложных по форме структур, как клетки, используется параметр, который называется эквивалентный радиус - радиус круга с той же площадью, что и у исследуемого объекта. Итоговая форма получается при помощи кубической сплайн-интерполяции, которая дает гладкий и эластичный контур. Аналогично можно генерировать форму ядра, цитоплазмы и других компонент клетки.

Не менее важным фактором при моделировании является текстура объектов. В данной модели используется вариант, в котором текстура в точке  $(x,y)$  задается как

$$t(x, y) = B + \sum_{i=0}^{n-1} p^i \eta_{xy}(2^i) \quad (2)$$

где вычисляется взвешенная сумма  $n$  октав функции шума  $\eta_{xy}$  с заданной частотой. Уровень шума задается степенью выгорания  $p$  и интенсивности константы  $B$ .

## IV. МОДЕЛИРОВАНИЕ ГРУПП КЛЕТОК

После определения свойств клеток следует перейти к рассмотрению параметров всего моделируемого изображения. И в первую очередь это касается количества и расположения клеток. Клетки в силу разнообразных биологических причин зачастую объединяются в кластеры. Заданное количество кластеров  $N_c$  случайно распределяется в изображении с координатами  $(x_c, y_c)$ . Клетки, отнесенные к

кластеру, располагаются вокруг центра кластера в соответствии с нормальным распределением и заданной вероятностью  $p_s$ , а с вероятностью  $(1-p_s)$  – равномерно по всему изображению.

В реальных изображениях клеточных популяций клетки часто плотно пространственно распределены или перекрываются, что может доставлять определенные трудности при их моделировании. Поэтому важно учесть возможность перекрытия и при моделировании изображения. Степень перекрытия определяется как

$$L_{ij} = \frac{|R_i \cap R_j|}{|R_j|}, i \neq j, \quad (3)$$

где  $R_i, R_j$  – области пикселей, занимаемые  $i$ -ой и  $j$ -ой клетками. Если же уровень перекрытия превышает максимально заданное значение, то в таком случае определяются новые координаты для  $j$ -ой клетки.

Завершающей стадией процесса моделирования является наложение на сгенерированное идеальное изображение погрешностей, наблюдаемых в реальной измерительной системе. Неравномерность освещения микроскопа, вызванную смещением источника свет и представляющую сложность при сегментации [3], можно смоделировать при помощи параболического полинома второй степени, а эффект автофлуоресценции с энергией  $E_a$  можно создать, медленно изменяя случайную текстуру.

Определенные погрешности вызывает трехмерная структура исследуемого образца ткани: в связи с тем, что только плоские объекты в фокальной плоскости будут находиться в фокусе, некоторые объекты будут получаться размытыми. Для моделирования эффекта размытия определена функция разброса точек (PSF, Point Spread Function) для каждого пикселя – с ее помощью импульс в точке  $(p, q)$  отображается в соседнюю точку  $(x, y)$  [4]. Для пространственного размытия аппроксимируем эту функцию распределением Гаусса [2].

Наконец, необходимо учесть влияние на уровень сигнала шума ФЭУ [5]. Этот шум можно аппроксимировать шумом Гаусса с нулевой медианой и дисперсией  $\sigma_d^2$ , которая и используется в данной модели вместо параметра отношения сигнал-шум [2], а сам шум обозначается  $N_g$ . Таким образом, идеальное изображение  $I$  с учетом неравномерным освещением и автофлуоресценцией фона примет вид:

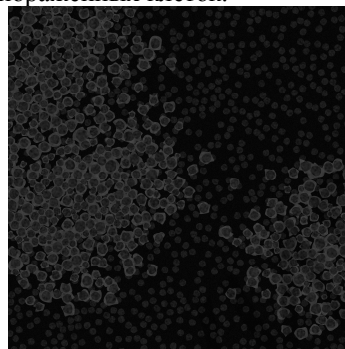
$$\hat{I} = \sum_{p, q \in R} PSF(x, y; p, q) I(p, q) + N_g, \quad (4)$$

где  $R$  – набор пикселей вокруг искаженной точки.

Измерительная система, как правило, записывает полученное изображение в файл с определенной степенью сжатия, что вызвано большими размерами исходного изображения. Ошибки сжатия зависят от используемого алгоритма сжатия, но, наиболее популярным является JPEG-сжатие, артефакты которого также учитываются в модели.

Приведенный алгоритм позволяет моделировать лишь одну популяцию клеток, в то время как экспериментальные изображения содержат две клеточные популяции: здоровые клетки и пораженные. В связи с этим необходимо последовательно генерировать два синтетических изображения здоровых и пораженных клеток опухоли при одних и тех же условиях проведения

эксперимента. Затем полученные изображения совмещаются вместе: при наложении изображений предпочтение отдается популяции пораженных клеток как объекту исследования. Таким образом, объекты популяции здоровых клеток будут присутствовать на итоговом изображении только в тех позициях, в которых не будет находиться ни один объект популяции пораженных клеток.



Смоделированное изображение. Представлены два кластера, включающие 500 раковых клеток. Также видны более тусклые здоровые клетки меньшего размера.

В ходе моделирования получены изображения, представляющие собой совокупность пораженных и здоровых клеток, что и демонстрирует рисунок 1. Проведенный анализ гистограмм распределения интенсивностей синтетических и экспериментальных изображений в различных цветовых каналах, а также сравнение их эквивалентных радиусов показывает близость реальных и смоделированных изображений, что указывает на адекватность разработанной модели.

## V. ЗАКЛЮЧЕНИЕ

В работе представлена и исследована имитационная модель для генерации синтетических изображений раковых опухолей. Задаваемые параметры моделирования позволяют получить изображения для многочисленных условий проведения эксперимента и различного состояния и структуры исследуемого биологического образца, что делает данную имитационную модель полезным инструментом при моделировании реально проводимых экспериментов. Следует отметить, что используемый подход при совмещении генерируемых изображений популяций клеток ведет к некоторой избыточности, т.к. часть сгенерированных объектов полностью перекрыта на полученном изображении, а также увеличивает вычислительные затраты, еще более продлевая и так весьма не быстрый процесс воспроизведения итогового изображения.

- [1] Карнаухов, В.Н. Люминесцентный анализ клеток. Пушкино, 2002.
- [2] A. Lehmussola, P. Ruusuvauro, J. Selinummi, H. Huttunen, O. Yli-Harja, "Computational framework for simulating fluorescence microscope images with cell populations". 2007, IEEE transactions on medical imaging, pp. 1010-1016.
- [3] Y. Lisitsa, M. Yatskou, V. Apanasovich, H. Rui, T. Apanasovich. Fully-automated segmentation of tumor nuclei in cancer tissue images.
- [4] Р. Гонсалес, Р. Вудс, С. Эддингс. Цифровая обработка изображений в среде MATLAB. Москва : Техносфера, 2006.
- [5] Hamamatsu Photonics K.K., Electron Tube Division. Photomultiplier tubes. Basics and applications. 2006.