
МЕТОД ОБУЧЕНИЯ КОЛИЧЕСТВЕННОМУ АНАЛИЗУ СОДЕРЖАНИЯ МИКРОСКОПИЧЕСКИХ ВИДЕОИЗОБРАЖЕНИЙ КАК ИНСТРУМЕНТ БИОИНФОРМАТИКИ

М.Ю. Чернышов

Президиум Иркутского научного центра СО РАН
ул. Лермонтова, 134, Иркутск, Россия, 664033

Н.В. Абасов

Институт систем энергетики им. Л.А. Мелентьева СО РАН
ул. Лермонтова, 130, Иркутск, Россия, 664033

В.Н. Нурминский, А.С. Столбиков

Национальный исследовательский Иркутский
государственный технический университет
ул. Лермонтова, 83, Иркутск, Россия, 664074

Предложены метод, аппаратный и программный комплексы, обеспечивающие эффективную поддержку процесса обучения студентов и профессионалов приемам количественного анализа содержания микроскопических видеоизображений клеточных препаратов в медико-биологических исследованиях. Программная реализация этого метода предполагает применение аналитических процедур фильтрации, сегментации видеоизображения, вычитания фона, бинаризации, а также аналитической процедуры подсчета объектов на микроизображении.

Ключевые слова: микроскопические медико-биологические препараты, медико-биологическая морфометрия, количественный анализ содержания микроскопических видеоизображений, метод количественного анализа.

Продиктованная требованиями времени постановка задачи повышения уровня и качества образования в современных вузах потребовала многоаспектной информатизации образования. Сегодня никого не удивят системы компьютеризированного контроля обучения, контроля успеваемости, системы ведения отчетности преподавателей и т.п. Реакцией на поставленную задачу стало не только повсеместное введение в вузах курсов информационных технологий (причем даже в вузах филологического и социологического профиля), но также открытие в ведущих вузах страны новых факультетов и отделений: факультетов информатизации (РГЭУ; ЮФУ), факультетов биоинформатики (МГУ), отделения биоинформатики (Первый Московский государственный медицинский университет им. И.М. Сеченова; НИ ИргТУ) и т.п.

Однако даже просто решение задач обучения на таких факультетах и проведение научных исследований студентами под руководством преподавателей требует разработки и использования в практике обучения новейших методов и технологий, которые, если следовать велениям времени, должны быть не просто информационными (термин «информационные» сам по себе несодержателен), а прежде всего телекоммуникационными или даже интеллектными (термин академика РАН С.Н. Васильева), т.е. предполагающими использование элементов искусственного интеллекта. Ведь совершенная обучающая система должна быть в равной мере

пригодна для обучения как студентов, так и специалистов, например, операторов, которым предстоит управлять движением и стыковкой космических аппаратов, или будущих специалистов в области биоинформатики.

Результатом предшествующего этапа работы группы исследователей под руководством М.Ю. Чернышова стала информационная технология содержательного (логико-смыслового) анализа программ [3] и построенная на ее основе система, предназначенная для информатизации обучения будущих программистов и квалифицированных пользователей [4; 5]. Новой разработкой группы стали метод, а также аппаратный и программный комплексы, построенные на его основе и предназначенные для медико-биологической морфометрии (т.е. структурного, содержательного анализа состава микрообъектов на изображении и подсчета объектов на нем). Они дополняют известные методики медико-биологической морфометрии. Программный комплекс, рассчитанный на повышение работоспособности метода, обеспечивает автоматизацию количественного анализа последовательности микроскопических видеоизображений в микрокуветах, т.е. анализ не отдельных кадров, а последовательности видеоизображений, получаемых при мониторинге биологических процессов, наблюдаемых под электронным микроскопом *in vitro* [6].

Практическое использование комплекса предполагает подготовку специалистов, которые могут быть задействованы в его установке, наладке и эксплуатации и могут включиться в исследования, поэтому на первом этапе на основе имеющегося комплекса были разработаны:

— метод обучения количественному анализу микроскопических видеоизображений, который представлен ниже;

— программный интерфейс, предназначенный для поддержки обучения приемам количественного анализа микроскопических видеоизображений.

Важным является их наличие в деле подготовки будущих специалистов в области биоинформатики. Созданные метод обучения, обучающий комплекс и программный интерфейс, реализующий основные функции обучения, ценны тем, что сегодня на рынке нет ни отечественных, ни зарубежных аппаратных и программных систем, с помощью которых можно было бы обучать технологии выполнения количественного анализа микроскопических изображений и обучать специалистов.

Основные функции аппаратного комплекса. Согласно технологии исследуется последовательность видеоизображений проб микрообъектов анализируемого препарата, находящихся в кюветах. Эти изображения регистрируются циклически с помощью установки цейтраферной компьютерной видеосъемки (УЦКВ) [2]. Установка имеет 12 микрокамер, в которые помещаются пробы. Размер кадра регистрации пробы составляет 1820×1365 мкм, что соответствует размеру единичного изображения в 1600×1200 пикселей. Последовательность видеоизображений проб в 12 микрокамерах подвергается анализу. Анализ выполняется автоматически с помощью программного комплекса.

Основные функции программного комплекса, предназначенного для количественного анализа микроскопических видеоизображений, и интерфейса к нему. Качественный и количественный анализ — важнейшие составляющие процесса исследования микроскопических видеоизображений. Выполнять эти виды

анализа вручную невозможно, необходимы новейшие средства биоинформатики, современные эффективные программные комплексы. В связи с этим важное значение в деле подготовки студентов-биологов и биоинформатиков приобретает обучение приемам анализа на таких комплексах. Обучение предполагает наличие интерфейса к имеющемуся аналитическому комплексу, позволяющего использовать возможности комплекса эффективно.

С точки зрения обучения важно, чтобы при работе с комплексом обучаемый студент-биолог мог:

— наблюдать исследуемое растровое видеоизображение препарата на всех фазах его анализа, выполняя качественный анализ содержания видеоизображения (объектов и отношений между ними);

— оценивать необходимость и перспективы количественного анализа объектов на видеоизображении (количественный анализ может потребовать преобразования последнего);

— осуществлять декомпозицию видеоизображения (что могло бы упростить его дальнейший анализ за счет уменьшения количества объектов изображения и связей между ними на каждом конкретном этапе анализа видеоизображения или его части);

— выполнять обоснованный количественный анализ состава видеоизображения (например числа объектов на нем или на его части);

— представлять результаты анализа в количественной форме.

Усилия коллектива предполагали создание такого программного обеспечения.

Обучение технологии количественного анализа. Обучение приемам получения количественной информации об объектах на видеоизображении складывается из следующих этапов:

- 1) усвоение метода анализа;
- 2) обучение работе на УЦКВ;
- 3) обучение работе с программным комплексом.

Содержание метода количественного анализа видеоизображений. Согласно традиционной практике до выполнения операций морфометрии объекты в медико-биологических препаратах выделяют, например, путем окрашивания с помощью специальных зондов-красителей. За счет окрашивания достигается цвето-яркостное различие объектов на окружающем фоне, регистрируемое непосредственно или при облучении препарата лазером в полосе длин волн λ от 10^{-5} до 1 нм. Эта задача решается проще, если препарат содержит микрообъекты, которые близки по размеру и с точки зрения восприятия. При этом морфометрия предполагает последовательное применение алгоритма выделения поля микрообъектов на растровом видеоизображении и алгоритма определения их количества. Компьютерная обработка данных морфометрии предполагает определения класса объектов на изображении и применение многошагового алгоритма:

- 1) предварительная обработка видеоизображения;
- 2) яркостная фильтрация;
- 3) вычитание фона;
- 4) бинаризация;
- 5) подсчет числа объектов на изображении.

Многошаговый алгоритм количественного анализа видеоизображений может быть применен в количественном анализе последовательностей изображений микрообъектов анализируемого препарата.

Алгоритм предварительной обработки видеоизображения предполагает улучшение яркости и контраста, т.е. различимости микрообъектов. Пусть $x(i, j) = x_{i,j}$ и $y(i, j) = y_{i,j}$ — соответственно, значения яркости исходного изображения и изображения, получаемого на выходе после обработки, в пикселе кадра, имеющем декартовы координаты i, j , которые могут сохраняться в матрице как номера строк i и столбцов j .

Поэлементная обработка изображения подразумевает, что строится функциональная зависимость между значениями яркости такая, что $y(i, j) = f_{i,j}(x_{i,j})$, которая по значению исходного сигнала позволяет определить значение выходного сигнала. Очевидно, что значения функции $f_{i,j}(x_{i,j})$ зависят от координат, т.е. в общем случае обработка кадра видеоизображения может быть неоднородной. Для простоты в наших расчетах мы предположили, что указанная зависимость между яркостями исходного и обработанного изображений является однородной, т.е. описывается функцией, значения которой одинаковы во всех точках кадра:

$$y = kx(x_{\max} - x),$$

где x_{\max} — максимальное значение исходного сигнала; k — константа, позволяющая управлять динамическим диапазоном преобразования изображения.

При $x = x_{\max}$ яркость обнуляется.

Задачу такой предварительной обработки видеоизображения можно решить, например, с помощью программы Adobe Photoshop. Однако она не дает удовлетворительной декомпозиции изображения. Кроме того, микрообъекты имеют сферическую форму, и необходима фокусировка на «срезах микрообъектов по середине» (а не в их верхней или нижней части).

Фильтрация видеоизображения выполняется в целях улучшения контраста и яркости, т.е. различимости микрообъектов (рис. 1).

Одна из проблем может быть связана с определением порога яркости m_{th} , с которым сравнивается яркость исходного видеоизображения и, таким образом, определяются значения яркости в каждой точке выходного изображения. Эффективным способом фильтрации изображения по яркости является Винеровская фильтрация. Подробное описание решения задачи фильтрации изображений с использованием оптимизационного подхода и решением уравнения Винера—Хопфа дано в статье М.С. Кириченко и С.В. Панина [1]. Решение этой задачи позволяет оценить результат линейной фильтрации полезного сигнала:

$$x^*(i, j) = \sum_{(i_1, j_1) \in S} \alpha(i_1, j_1) y(i - i_1, j - j_1),$$

где $x^*(i, j) = x^*_{i,j}$ — результат фильтрации полезного сигнала $x(i, j) = x_{i,j}$ в точке видеоизображения с координатами (i, j) ; S — множество координат точек, образующих окрестность точки изображения; $\alpha(i_1, j_1)$ — весовые коэффициенты, совокупность которых дает двумерную импульсную характеристику фильтра.

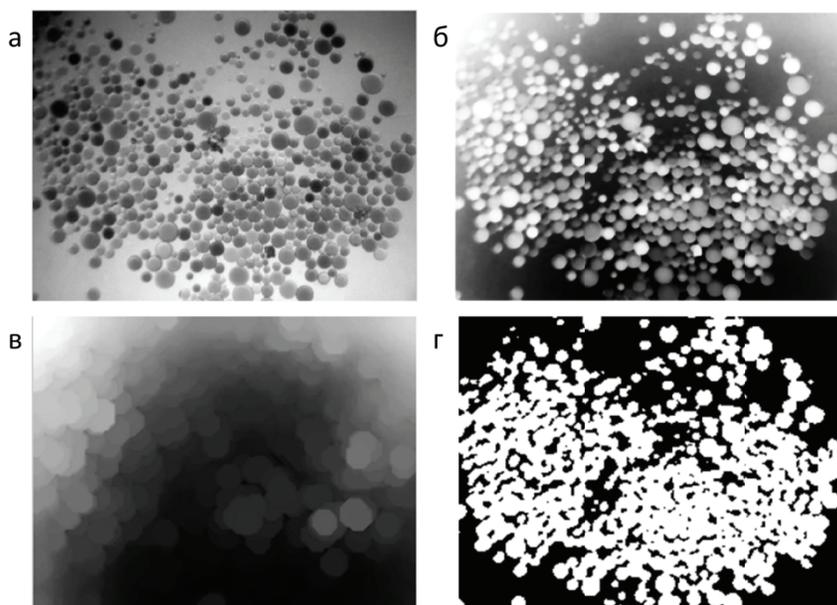


Рис. 1. Алгоритм выделения группы клеток для их подсчета:

а — исходное изображение; б — итог первичной обработки и фильтрации;
в — вычитаемый фон; г — результат бинаризации

В свою очередь, это позволяет, например, оценить изменение средней яркости видеоизображения при его линейной фильтрации путем вычисления математического ожидания результата фильтрации полезного сигнала:

$$m_y^*(i, j) = m_x \sum_{(i, j) \in S} \sum \alpha(i_1, j_1),$$

где предполагается, что средняя яркость m_x входного изображения $x(\cdot)$ не зависит от координат, и, как следствие, средняя яркость m_x^* изображения $y^*(\cdot)$ на выходе фильтра тоже постоянна во всех точках кадра. В ходе анализа важно поддерживать равномерную среднюю яркость кадра. Это достигается при выполнении равенства

$$\sum_{(i, j) \in S} \sum \alpha(i_1, j_1) = 1.$$

Задача вычитания фона и бинаризации видеоизображения для выделения объектов по морфологическим (текстурным) особенностям может быть решена путем его представления на языке теории множеств и соответствующих преобразований. Ее решение проще, если можно избежать сравнительного анализа цветов (объекты однородны по цвету). Тогда полутоновое изображение преобразуется в бинарное, и это все упрощает.

Поясним содержание процедуры бинаризации. Представим распределение яркости полутонового видеоизображения функцией $f(x, y)$, принимающей только неотрицательные значения, и предположим для простоты, что она имеет целые и равномерно разнесенные амплитуды яркости $m = 0, 1, 2, \dots$. Эта функция может быть реинтерпретирована как сумма бинарных функций $f_b(i, j)$, принимающих

только бинарные значения, соответствующие исходному полутоновому видеоизображению:

$$f(x, y) = \sum_{(i,j) \in S} f_b(i, j),$$

$$\text{где } f_b(i, j) = \begin{cases} 1 & \text{при } f(x, y) \geq m_{th} \\ 0 & \text{при } f(x, y) < m_{th} \end{cases}.$$

Иными словами, каждая бинарная функция $f_b(i, j)$ может быть представлена с помощью пороговых множеств X_b на плоскости, причем для каждой точки бинарного изображения: $(i, j) \in X_b$, если $f_b(i, j) = 1$, и $(i, j) \notin X_b$, если $f_b(i, j) = 0$.

В результате этого преобразования растровое полутоновое изображение преобразуется в бинарное (см. рис. 1г), которое соответствует матрице, содержащей только нули и единицы. Оно может быть сохранено в виде файла формата .txt, который затем используется на этапе количественного анализа, реализуемого с помощью специальной подпрограммы, вычисляющей относительную площадь изображения, занятую микрообъектами. Для решения задачи бинаризации может быть использована, например, система Mathcad 2000, преобразующая электронные изображения формата .bmp в матрицу, состоящую из чисел от 0 до 255, соответствующих пикселям монохромного видеоизображения. Пороговым ограничением m_{th} амплитуды яркости является некоторое число из указанного диапазона. В нашем случае значение этого порога было подобрано экспериментально и составило $m_{th} = 24$.

Задача подсчета числа объектов на видеоизображении. Решение этой задачи существенно упрощается, если микрообъекты имеют равные или близкие диаметры. Тогда достаточно определить площадь (в пикселях) области, занятой микрообъектами, и поделить результат на диаметр объекта.

Определение количества M микрообъектов, идентифицируемых в полезном поле видеоизображения, осуществляется в два шага: 1) площадь полезного поля S_u определяется как количество пикселей n переднего плана (т.е. тех, которым соответствует значение $f_b = 1$) на общем поле видеоизображения, содержащем N пикселей, т.е. $S_u = n = \eta N$; 2) S_u делится на структурообразующий параметр δ , соответствующий размеру одного анализируемого микрообъекта в пикселях: $M = \eta N / \delta$.

Программная реализация метода. Данный метод реализован пока по старой технологии как программный продукт в средах MATLAB (программа CellsPro) и Visual Studio C++ с использованием процедур обработки изображений библиотеки OpenCV (программа CellsProCV) [2].

Разработанные нами метод, аппаратный комплекс и программный комплекс, предназначенные для медико-биологической морфометрии (т.е. содержательного, структурного анализа состава микрообъектов на изображении и подсчета объектов

на нем) были подвергнуты тестированию, результаты которого отражены в [3] (при изменении цвето-яркостных характеристик изображения на 20—50% вариабельность количественных результатов — не более 5%). Таким образом, они вносят вклад в развитие средств медико-биологической морфометрии и создают основу для нового метода качественной медицинской диагностики заболеваний.

Судя по характеристикам, созданный программный комплекс обеспечивает автоматизацию количественного анализа последовательности микроскопических видеоизображений, может повысить эффективность мониторинга биологических процессов, наблюдаемых под микроскопом *in vitro*. На основе этого комплекса создан обучающий комплекс, обеспечивающий выполнение наглядно контролируемых пользователем аналитических операций и, таким образом, позволяющий обучать студентов и профессионалов-аналитиков высокотехнологичным приемам количественного анализа состава микрообъектов.

Важно уже то, что в условиях дефицита инновационных средств, необходимых для подготовки специалистов в области биоинформатики в вузах, созданный комплекс может стать первым таким средством. Он ценен уже тем, что сегодня на рынке нет ни отечественных, ни зарубежных программных комплексов, на которых можно было бы обучать таких специалистов. Важным достоинством созданного программного комплекса является использование в нем элементов новой технологии вычислительного моделирования, предложенной в [3; 4]. Поэтому в перспективе возможности комплекса не ограничены лишь простым морфологическим анализом.

Усовершенствованный программный комплекс, создаваемый в настоящее время на основе прежнего, должен позволить выполнять элементы логико-смыслового анализа содержания видеоизображений. Он позволит не только быстро и качественно обучать приемам анализа, но также оперативно автоматически обрабатывать результаты анализа и строить аналитические зависимости, которые могут иметь различные формы представления. Это весьма удобно при обучении студентов и квалифицированных аналитиков.

ЛИТЕРАТУРА

- [1] *Кириченко М.С., Панин С.В.* Разработка адаптивного алгоритма оценки информативности динамических признаков для обработки и анализа изображений // *Вычислительные технологии*. — 2005. — Т. 10. — № 1. — С. 58—70.
- [2] *Нурминский В.Н., Озолина Н.В., Нестеркина И.С., Колесникова Е.В. и др.* Стабильность вакуолярных мембран растений при осмотическом стрессе и воздействии редокс-агентов // *Биологические мембраны*. — 2011. — Т. 28. — № 3. — С. 224—229.
- [3] *Чернышов М.Ю., Абасов Н.В.* Основы вычислительной технологии, предназначенной для исследования программных систем и основанной на принципах логико-смыслового анализа // *Вестник ТОГУ. Серия «Информатика, вычислительная техника, управление»*. — 2012. — № 2 (25). — С. 27—36.
- [4] *Чернышов М.Ю.* Система содержательного анализа текстов программ, предназначенная для обучения программистов аналитическим приемам // *Аналитическая механика, устойчивость, управление: Материалы X Междун. Четаевской конференции*. — Казань, 2012.

- [5] Чернышов М.Ю. Система содержательного анализа программ, предназначенная для информатизации обучения будущих программистов и квалифицированных пользователей // Вестник Российского университета дружбы народов. Серия «Информатизация образования». — 2012. — № 2. — С. 111—117.
- [6] Чернышов М.Ю., Нурминский В.Н., Абасов Н.В., Осипчук Е.Н. Метод количественного анализа объектов на видеоизображениях, рассчитанный на использование в ГИС-технологиях и анализе микроскопических видеоизображений // Вестник Бурятского государственного университета. Серия «Математика. Информатика». — 2012. — Вып. 9. — С. 76—80.

A METHOD FOR TRAINING IN QUANTITATIVE ANALYSIS OF THE CONTENT OF MICROSCOPIC VIDEO-IMAGES AS A TOOL OF BIOINFORMATICS

M. Yu. Chernyshov

Presidium of Irkutsk Scientific Center, Siberian Branch of RAS
Lermontov str., 134, Irkutsk, Russia, 664033

N.V. Abasov

Melentiev Energy System Institute, Siberian Branch of RAS
Lermontov str., 130, Irkutsk, Russia, 664033

V.N. Nurminsky, A.S. Stolbikov

National Research Irkutsk State Technical University
Lermontov str., 83, Irkutsk, Russia, 664074

A method, a hardware and a software complexes, which provide for efficient support of the process of teaching students and professionals in techniques of quantitative analysis of the content of microscopic images of cell preparations in medical and biological investigations, are proposed. The software implementation of this method presumes application of analytical procedures of filtering, segmentation of the video-image, subtracting the background, binarization and the analytical procedure of computing objects on the microscopic image.

Key words: microscopic images of medical and biological preparations, medical and biological morphometric analysis, quantitative analysis of the content of microscopic images; technique of quantitative analysis.