

О.Н. Березский, Г.Н. Мельник

Информационная технология анализа и синтеза гистологических изображений в системах автоматизированной микроскопии

Разработана информационная технология анализа и синтеза гистологических изображений, в основе которой лежат структурные методы анализа и синтеза изображений на базе плоских кристаллографических групп. Описан программный комплекс – система автоматизированной микроскопии, реализующая предложенную информационную технологию.

Information technology for analysis and synthesis of histological images is developed. Information technology is based on structural methods of analysis and synthesis of images based on plane crystallographic groups. The automated microscopy software system – is described. It implements the proposed information technology.

Розроблено інформаційну технологію аналізу та синтезу гістологічних зображень, заснованих на структурних методах аналізу та синтезу зображень на базі плоских кристаллографічних груп. Описано програмний комплекс – систему автоматизованої мікроскопії, що реалізує запропоновану інформаційну технологію.

Введение. В клинической практике при исследовании злокачественных новообразований для постановки диагноза используется световая микроскопия – источник гистологических изображений (ГИ) и область лабораторной диагностики, где доминирует трудоемкий субъективный качественный анализ.

Рак – одна из наиболее распространенных причин смерти на планете, и по данным Всемирной организации здравоохранения до 2030 года глобальная смертность от него в мире в сравнении с 2007 годом вырастет в полтора раза и составит 11,5 миллионов случаев. Диагностирование раковых заболеваний требует гистопатологического анализа, т.е. исследования поврежденной ткани человека.

Для анализа ГИ на рынке существует широкий ряд программно-аппаратных комплексов – систем автоматизированной микроскопии (САМ). Они отличаются по функциональным возможностям, специализации, архитектуре, уровню автоматизации процесса анализа.

Согласно функциональным возможностям САМ можно разделить на такие типы:

- системы получения изображения из фото- и видеокамеры или цифрового микроскопа;
- системы, содержащие средства для ручного интерактивного измерения размеров микрообъектов и структур, подсчета их количества и статистического анализа;

- системы, которые имеют средства для автоматизированного выделения микрообъектов путем подбора параметров операции сегментации;
- системы, содержащие средства классификации изображения, микрообъектов или состояний микрообъекта;
- экспертные системы, позволяющие строить качественное описание состояний микрообъектов.

САМ реализуют информационную технологию [1, 2] анализа гистологических изображений. Исходя из типа решаемой задачи и использованных операций, программное обеспечение САМ можно условно разделить на три класса (табл. 1): редакторы изображений, автоматизированные системы обработки изображений [3–7], экспертные системы анализа изображений [8–11].

Обобщенная структура САМ [12–15] приведена на рис. 1. В состав САМ входит осветительно-фокусирующее устройство, гистологический препарат, микроскоп, цифровая камера, блок информационной технологии и врач-диагност.

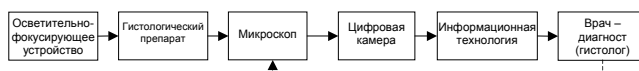


Рис. 1. Обобщенная структурная схема САМ

При проведении морфометрического исследования [16–24] врач-диагност (гистолог) выполняет операции поиска нужного объекта, идентификации его границ и вычисление числовых признаков. Выделим следующие базовые операции морфометрического анализа:

Ключевые слова: гистологическое изображение, анализ, синтез, система автоматизированной микроскопии, информационная технология.

- создание виртуального слайда;
- предварительная обработка;
- расширение глубины фокуса;
- сегментация по интенсивности;
- текстурная сегментация;
- выделение контуров;
- детекция ядер клеток, цитоплазмы, слоев;
- оценка структурных изменений.

Существующие системы автоматизируют от четырех до шести перечисленных операций, тогда как детекция слоев выполняется путем ручного интерактивного выделения, а оценка структурных изменений проводится путем качественной визуальной оценки.

Т а б л и ц а 1. Классификация программных средств САМ

Класс	Редактор изображений	Автоматизированная система обработки изображений	Экспертная система анализа изображений
Задачи	Получение, предварительная обработка изображения	Ручное и автоматическое выделение микрообъектов, подсчет их количества, измерение числовых признаков, статистическая обработка результатов	Логический вывод о типе злокачественного процесса
Операции	Управление микроскопом и камерой, нормализация фона, удаление шумов, подчеркивание границ, улучшение контраста	Сегментация, контурный анализ, текстурный анализ, детекция микрообъектов, классификация изображений	Распознавание изображений, логический вывод на основе нечеткой логики, логический вывод на основе базы знаний
Пример существующих систем	ScreenMeter, Motic Images Advanced 3.2	ВидеоТесТ-Морфология 5.0, AnalySIS Five, BioVision, QCapture PRO 6.0, CellProfiler	KIDS

Постановка задачи

Проведенный анализ существующих САМ показывает, что для части диагностических процессов, базирующихся на структурных характеристиках тканей и взаимном расположении клеток, математическое и алгоритмическое обеспечение отсутствует. Потому актуальна задача – разработка методов и алгоритмов структурного анализа и синтеза гистологических изображений и построение на их основе информационной технологии.

Метод структурного анализа гистологических изображений

Для анализа структурных патологических изменений в ткани разработан метод и алгоритмы анализа ГИ на основе теории кристаллографической группы [25]. Анализ ГИ состоит из следующих этапов:

- определение соответствующих точек на контурах микрообъектов;
- вычисление аффинных преобразований между микрообъектами;
- определение группы симметрии для группы микрообъектов (части изображения).

Соответствующие точки определяются на основе дискретных центральных моментов области: вычисляется угол наклона главной оси, координаты центра инерции объекта и т.д.. Для области выделяются три соответствующие контурные точки, на основе которых рассчитываются коэффициенты аффинных преобразований.

ГИ представлено как асимметричное изображение. Определение асимметричного изображения дано через симметричное [26], состоящее из элементарного изображения (клетки) I_E , на основе которого, с помощью применения оператора геометрических преобразований S , получен раппорт:

$$P = S(I_E). \quad (1)$$

Тогда уравнение асимметричного изображения выглядит так:

$$I_A = jL_Y D_Y (iL_X D_X \times \times [T_n D_n (T_{n-1} D_{n-1} (T_{n-2} D_{n-2} \dots T_1 D_1 X))]), \quad (2)$$

где X – матрица координат элементарного изображения; T_1, T_2, \dots, T_n – матрицы порождающих преобразований раппорта; D_1, D_2, \dots, D_n – матрицы искажений раппорта; L_X, L_Y – матрицы трансляции вдоль OX и OY ; D_X, D_Y – матрицы искажений оси трансляции относительно OX и OY ; i, j – количество итерации сдвига раппорта вдоль оси OX и OY соответственно.

Скомпоновав порождающие преобразования T и искажение D в комбинированное преобразование C , имеем:

$$I_A = jC_Y (iC_X [C_n (C_{n-1} (C_{n-2} \dots C_1 X))]).$$

Для оценки искажений структуры микро-объектов на ГИ нужно найти матрицы T и D .

Для анализа ГИ предложен метод [27], заключающийся в представлении его в виде дерева аффинных преобразований (ДАП):

$$GR=(V, A),$$

где V – множество вершин, представляющих элементарные рисунки; A – множество ветвей.

Множество ветвей A состоит из множества ветвей видов преобразований симметрии A_C и множества ветвей сдвигов A_L :

$$A = A_C \cup A_L.$$

Последовательность ветвей ДАП для множества k элементарных изображений представим в виде списка пар значений (C, L) :

$$g(A) = \{(C_i, L_j) | i = 1..k, j = 1..k\}.$$

Группу симметрии раппорта опишем через упорядоченную последовательность порождающих преобразований между первым элементарным изображением и всеми другими:

$$g^*(P) = \{(T_i, L_j) | T \in T^*, i = 1..k, j = 1..k\},$$

где T^* – множество всех порождающих преобразований симметрии для 17 групп на плоскости, L – множество всех матриц сдвига, которые могут образовать группы.

Определение группы симметрии раппорта сводится к последовательному поэлементному сравнению найденного списка $g(A)$ со списками порождающих преобразований каждой из 17 групп $g_r^*(P)$, $r = 1, \dots, 17$.

Коэффициент искажения для одного порождающего преобразования равен:

$$d_u = \sum_{i=1}^2 \sum_{j=1}^2 |c_{ij} - t_{ij}^u|,$$

где c_{ij} – элемент матрицы C , t_{ij}^u – элемент матрицы T_u , $u = \overline{1, k}$, k – количество элементарных изображений в раппорте.

Для раппорта определяем общий коэффициент искажения:

$$\delta_p = \frac{\sum_{u=1}^k d_u}{k}. \quad (3)$$

Для оценки структурных изменений ткани при злокачественных новообразованиях введем коэффициент структурной атипии для изображения:

$$\delta_l = \frac{\sum_{i=1}^n \delta_{p_i}}{n}, \quad (4)$$

где n – количество раппортов на изображении.

Проведено экспериментальное исследование коэффициента структурной атипии (4) путем классификации тренировочной выборки ГИ тканей эпителиального слоя с известным классом изображения, вычислены показатели чувствительности и специфичности (табл. 2). При этом использованы БД объемом 400 изображений.

Т а б л и ц а 2. Коэффициент структурной атипии ткани

Классы изображений	Коэффициент структурной атипии	Чувствительность	Специфичность
Тяжелая дисплазия	0,54–0,67	0,96	0,93
Умеренная дисплазия	0,42–0,49	0,93	0,92
Нормальная ткань	0,25–0,41	0,95	0,94

Результаты исследований подтвердили характерность разработанного показателя во время анализа эпителиальной ткани на ГИ.

Метод синтеза гистологических изображений

Для формирования обучающих выборок использованы синтез гистологических изображений на основе плоских кристаллографических групп. В качестве непроектируемого элемента (элементарного изображения) взято изображение клетки, над которым выполнены порождающие преобразования определенной группы симметрии. Для моделирования структурных изменений в тканях осуществлены параметрические искажения порождающих преобразований, в результате чего получен ряд структур: регулярные, близко-регулярные, нерегулярные, близко-стохастические и стохастические.

Метод синтеза ГИ базируется на обходе ДАП с модифицированным коэффициентом структурной атипии и состоит из следующих шагов:

1. Задать значение коэффициентов искажения (3) и (4) для найденного в результате анализа ДАП.

2. Задать параметры синтеза изображений микрообъектов.

3. Синтезировать микрообъект.

4. Для всех элементов раппорта повторить шаг 3 в соответствии с (1).

5. Для каждого раппорта выполнить трансляцию в соответствии с (2).

6. Растеризовать изображение.

Определен объем памяти, необходимый для хранения параметров синтеза одного ГИ (табл. 3).

Таблица 3. Объем памяти, необходимый для хранения ГИ

Количество микрообъектов	10	100	1000
Объем для сгенерированного изображения 1Мп, байт	3145728	3145728	3145728
Объем для хранения параметров синтеза	11708	23048	136448
Коэффициент уменьшения объема памяти	268,7	136,5	23,1

Результаты экспериментов показывают, что объем памяти для хранения тестовой выборки изображений в наихудшем случае уменьшен до 23 раз.

Информационная модель данных гистологических изображений

Разработанная информационная модель [25] данных ГИ показана на рис. 2,а. Интерфейс, обеспечивающий отображение информационной модели, приведен на рис. 2,б.

В информационной модели использованы следующие массивы: цветовой признак (ЦП), матрицы искажения (МИ), порождающие преобразования (ПП), векторы трансляции и искажения. Источниками исходных информационных массивов есть блоки: текстурной сегментации, выделения элементарных изображений и их признаков, определения матриц преобразования элементарных изображений, распознавания отдельных раппортов и классификация их группы симметрии. Алгоритм сегментации определяется в зависимости от характеристик и типа изображения. На сегментированном изображении выделяются объекты и фон, вычисляются признаки объектов, определяются соответствующие точки на контурах объектов. На основе координат соответствующих точек рассчитывают матрицы аффинных

преобразований. Для каждой найденной матрицы преобразований определяют матрицу порождающих преобразований и матрицу искажения. На основе матриц порождающих преобразований классифицируют группу симметрии раппорта.

Процесс синтеза гистологического изображения начинается с синтеза элементарного изображения. Используя элементарное изображение, преобразование симметрии соответствующей группы и значение искажения генерируем раппорт. Построение гистологических изображений на плоскости проходит в три этапа:

- синтез элементарного изображения;
- размещение элементарного изображения в соответствии с преобразованиями определенной группы симметрии и заданной матрицей искажения (построение раппорта);
- построение необходимого количества раппортов.

Синтез раппорта соответствующей группы заключается в последовательном приложении к элементарному изображению преобразований заданной группы симметрии.

В точке диалога врач–диагност получает информацию о размещении структур микрообъектов и коэффициенте их искажения (КИ). На ее основе он принимает решение об изменении параметров синтеза изображения микрообъектов (ПСМО) для формирования тестовой выборки ГИ.

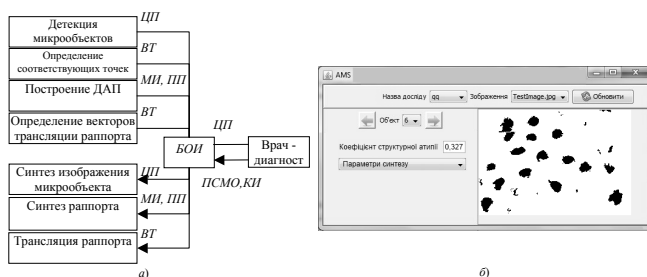


Рис. 2. Информационная модель данных гистологических изображений

Блок обработки информации (БОИ) предназначен для преобразования информации о границах слоев ткани, группы симметрии раппортов структур микрообъектов, коэффициентов искажения и обеспечения функции уп-

равления процессом синтеза ГИ. Задаваемыми врачом–диагностом параметрами синтеза являются границы структур микрообъектов, коэффициент искажения структур микрообъектов и количество слоев эпителиального слоя тканей.

Программная реализация САМ

Разработанное математическое и алгоритмическое обеспечение технологии обработки ГИ применено при создании программного обеспечения (ПО) САМ. Для написания программного средства использованы программная среда *IntelliJIDEA*, языки программирования *Java*, *Object Pascal*, *C++*. Для работы ПО необходима рабочая станция на базе операционной системы *Windows*. Программное средство разработано как расширение (модуль) редактора *ImageJ*.

Для проведения анализа входного изображения использована библиотека операций программного средства *ImageJ*. Разработанное ПО решает следующие задачи обработки ГИ: создание методики исследования (сценария), проведение исследования в автоматическом режиме, работа с БД, работа с правилами диагностирования. Модульная структура САМ изображена на рис. 3.



Рис. 3. Структура программного средства анализа и синтеза ГИ

Разработанные методы и алгоритмы программно реализованы в виде блоков вычислений количественных признаков структуры микрообъектов и синтеза ГИ.

Исследование уровня автоматизации САМ

Уровень автоматизации процесса проработки ГИ определен так:

$$P = \frac{K_A}{K} * 100\%,$$

где K_A – количество автоматических операций, K – общее количество операций.

Рассчитан уровень автоматизации обработки ГИ для самых распространенных систем (табл. 4). Ручные операции обозначены буквой *p*, автоматические – буквой *a*.

Таблица 4. Операции обработки ГИ

ПО САМ	Цена, у.е.	Предварительная обработка	Виртуальный слайд	Расшир. глубина фокуса	Сегментация по интенсивности	Сегментация по текстуре	Сегментация контурная	Детекция ядер клеток	Детекция цитоплазмы	Детекция слоев	Оценка структурных изменений	Уровень автоматизации, %
<i>Carl Zeiss Axio-Vision 4.8.2</i>	3000	p	a	a	a	p	a	p	p	p	p	40
<i>ImagePro Plus</i>	5500	p	a	a	a	p	p	p	p	p	p	30
<i>Cell Profiler</i>	0	p	a	a	a	a	a	p	p	p	p	50
Видеотест Морфология 5.2	4280	a	a	a	a	a	a	p	p	p	p	60
<i>ImageJ</i>	0	a	p	a	a	a	a	p	p	p	p	50
Разработанное ПО САМ	1500	a	p	a	a	a	a	p	p	a	a	70

Низкий уровень автоматизации обработки ГИ состоит в применении операций предварительной обработки и сегментации однородных областей на основе интенсивности или цвета пикселей. Средний уровень достигается при применении операций сегментации на основе текстурных признаков и детекции клеток. Высокий уровень автоматизации обеспечивается созданием виртуального слайда и расширения глубины фокуса. На рис. 4 изображена гистограмма уровня автоматизации систем аналогов.

Сравнительный анализ подтверждает эффективность предложенной информационной технологии, поскольку ее применение позволило увеличить уровень автоматизации не менее чем

на 10 процентов в сравнении с системой-аналогом Видеотест Морфология (Россия).

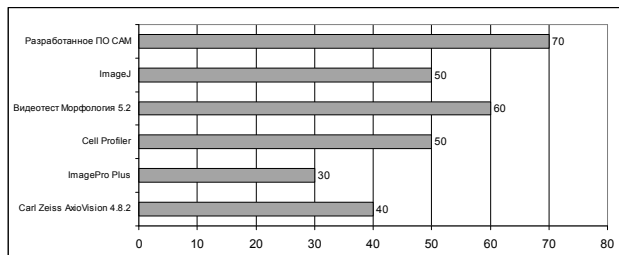


Рис. 4. Уровень автоматизации САМ

Проведены экспериментальные исследования оперативности обработки изображений эпителиального слоя шейки матки. Оценено среднее время выполнения (табл. 5) следующих операций: создание виртуального слайда, предварительной обработки, сегментации, детекции ядер клеток, детекция эпителиальных слоев, оценка структурных изменений.

Таблица 5. Среднее время выполнения операции при исследовании изображений эпителия шейки матки

Операция обработки и ее среднее время выполнения, с	Виртуальный слайд	Предварительная обработка	Сегментация по интенсивности	Детекция ядер клеток	Детекция слоев	Оценка структурных изменений	Суммарное время
САМ	t_1 , мин.	t_2 , мин.	t_3 , мин.	t_4 , мин.	t_5 , мин.	t_6 , мин.	T_c , мин.
Carl Zeiss Axio-Vision 4.8.2	1,4	1,1	0,7	0,8	3,8	3,7	11,5
ImagePro Plus	1,6	0,4	0,4	1,9	2,7	6,6	13,6
Cell Profiler	1	0,8	1,2	1,6	1,6	3,3	9,5
Видеотест Морфология 5.2	1,7	1	0,8	1,1	3,1	3,1	10,8
ImageJ	1,6	0,2	1,9	0,6	1,9	3,1	9,3
Разработанное ПО САМ	1,6	0,2	1,9	0,6	0,7	0,2	5,2

Для оценки оперативности при применении разработанной технологии среднее время выполнения совокупности операций определены как:

$$T_c = T_1 + T_i + \dots + T_n,$$

где T_i , ($i = 1, \dots, n$) – время выполнения i -й операции.

Результат сравнения среднего времени выполнения морфометрического анализа ГИ приведен на рис. 5.

Размер тестовой выборки составил 400 изображений. Применение разработанной техно-

логии дало возможность уменьшить среднее время обработки с 9,3 до 5,2 мин. Перспективное направление дальнейших исследований – использование разработанного коэффициента структурной атипичности как диагностического признака при изучении патологических процессов в различных органах человека.

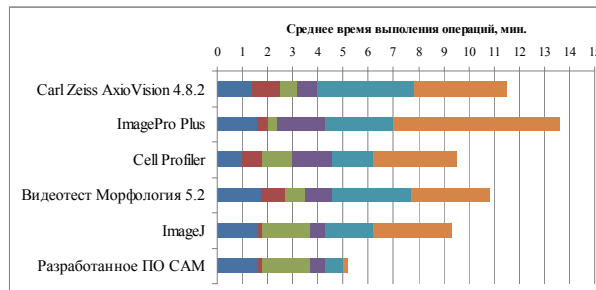


Рис. 5. Сравнительная оценка оперативности САМ

Вывод. Созданная информационная технология, реализованная в виде программного средства САМ, дала возможность повысить уровень автоматизации и оперативности обработки гистологических изображений.

Предложенные метод и алгоритмы синтеза ГИ значительно расширили базу тестовых изображений для проведения верификации алгоритмов распознавания изображений опухолевых тканей. Разработанные методы и алгоритмы анализа ГИ повысили достоверность диагностирования злокачественных новообразований путем количественной оценки патологических изменений в тканях человека.

1. Абламейко С.В., Лагуновский Д.М. Обработка изображений: технология, методы, применение. Учебное пособие. – Минск: Амалфея, 2000. – 304 с.
2. Інформаційні технології та моделювання бізнес-процесів. Навч. посібник / О.М. Томашевський, Г.Г. Цегелик, М.Б. Вітер та ін. – К.: Центр учбової літератури, 2012. – 296 с.
3. <http://rsbweb.nih.gov/ij/docs/index.html>
4. VideoTest – <http://www.videotest.ru/ru/products/4>
5. <http://www.cellprofiler.org/manuals.shtml>
6. <http://www.mediacy.com/index.aspx?page=IPP>
7. http://microscopy.zeiss.com/microscopy/en_de/products/microscope-software/axiovision-for-biology.html
8. Ovalle A., Garbay C. KIDS: A Distributed Expert System for Biomedical Image Interpretation. // Inform. Proc. in Medical Imaging. – 1991. – P. 419–433.
9. Ovalle A., Hugonnard E., Garbay C. Medical system design and knowledge acquisition using cooperating in-

- telligent agents. A case study for breast cancer diagnosis // *Artificial Intelligence in Medicine* – 1993. – N 10. – P. 247–258.
10. *Продеев А.Н., Захарова Е.Н.* Экспертные системы в медицине – К.: Век, 1998. – 320 с.
 11. *Пат. № 2293524* Россия, Способ дифференциальной диагностики фолликулярной аденомы и фолликулярного рака щитовидной железы / Т.Л. Полоз, А.В. Демин, В.А. Шкурупий, Опубл.: 02.20.2007. – 10 с.
 12. *Maintz J.B.A., Viergever M.A.* A Survey of Medical Image Registration // *Medical Image Analysis*. – 1998. – 1, N 2. – P. 1–37.
 13. *Bankman I.N.* Handbook of medical imaging // Academic Press, 2000. – 911 p.
 14. *Rangayyan R.M.* Biomedical Image Analysis. – USA: CRC PRESS, 2004. – 1312 p.
 15. *Microscopy Techniques. Advances in Biochemical Engineering/Biotechnology.* / Ed. by J. Rietdorf. – Berlin Heidelberg: Springer Verlag, 2005. – 322 p.
 16. *Claridge E., Cotton S., Moncrieff M.* From colour to tissue histology: physics based interpretation of images of pigmented skin lesions // *Medical Image Computing and Computer-Assisted Intervention* – Berlin: Springer, 2002. – P. 489–502.
 17. *Обработка изображений гистологических срезов и распознавание опухолевых клеток* / Ю.А. Скобцов, Т.В. Мартыненко, Ю.К. Гульков и др. // *Материалы II наук.-практ. конф. «Донбас-2020: Наука і техніка – виробництву»*, Донецьк 03–04 лют. 2004 р. – ДонНТУ, 2004.– С. 566–569.
 18. *Недзьведь А.М., Фридман М.В., Папок В.Е.* Современные возможности обработки изображений при морфологическом исследовании карцином щитовидной железы // *Медицинские новости*. – 2006. – № 12. – С. 115–119
 19. *Large-scale computations on histology images reveal grade-differentiating parameters for breast cancer* / S. Petushi, F. Garcia, M. Haber et al. // *BMC Medical Imaging* – 2006. – 1, N 6. – P. 14.
 20. *Karacali B., Vamvakidou A., Tozeren A.* Automated recognition of cell phenotypes in histology images based on membrane- and nuclei-targeting biomarkers // *BMC Medical Imaging* – 2007. – N 7. – P. 7. – <http://www.Biomeduntra.com/1471-2342/7/7>
 21. *Towards efficient automated characterization of irregular histology images via transformation to frieze-like patterns* / B.A. Canada, G.K. Thomas, K.C. Cheng et al. // *CIVR'08: Proc. of the 2008 Int. Conf. on Content-based image and video retrieval*. – New York: ACM, 2008. – P. 581–590.
 22. *Characterization of chromatin texture by contour complexity for cancer cell classification* / T. Kiyuna, A. Saito, E. Kerr et al. // *8th IEEE Int. Conf. on BioInformatics and BioEngineering, 2008* – Athens. – P. 1–6.
 23. *Wang C.W.* Robust Automated Tumour Segmentation on Histological and Immunohistochemical Tissue Images // *PLoS ONE* – 2011. – 2, N 6. – P. 1–8.
 24. *Медовый В.С.* Современный уровень автоматизации методик микроскопического анализа // *Медицинский алфавит: Лабораторная диагностика* – 2007. – № 2. – С. 4–6.
 25. *Березький О.М., Мельник Г.М.* Інформаційна технологія аналізу та синтезу гістологічних зображень // *Оброблення сигналів і зображень та розпізнавання образів: Праці Одиннадцятої всеукр. міжнар. конф. 15–19 жовт. 2012 р.* (УкрОбраз) – К.: УкрІНТЕІ, 2012. – С. 161–165.
 26. *Березький О.Н.* Теоретико-групповой подход к анализу симметричных изображений // *УСиМ*. – 2010. – № 6. – С. 16–24.
 27. *Мельник Г.М.* Метод і алгоритми аналізу симетричних зображень // *Штучний інтелект* – 2010. – № 4. – С. 253–261.

Тел. для справок: +38 0352 435-910,
+38 050 377-4579, 050 244-5088 (Тернополь)
E-mail: ob@tneu.edu.ua, mgm@tneu.edu.ua
© О.Н. Березький, Г.Н. Мельник, 2013