

МЕТОД ЗНАХОДЖЕННЯ ВІДПОВІДНИХ ТОЧОК НА КОНТУРАХ МІКРООБ'ЄКТІВ БІОМЕДИЧНОЇ ПРИРОДИ

© Мельник Г., 2012

Характерною властивістю цифрових біомедичних зображень є наявність повторюваних елементів з нечіткою геометричною структурою. Наближення форми клітин до форми круга або еліпса вимагає вдосконалення алгоритмів співставлення областей. Для визначення найкращого афінного перетворення між областями використано симетричну різницю множин точок областей. Застосування дискретних центральних моментів області дозволило підвищити точність визначення відповідних точок на контурах при визначенні афінних перетворень між областями мікрооб'єктів.

Ключові слова: біомедичні зображення, структурний аналіз, співставлення областей, центральні моменти, контур, відповідні точки.

The characteristic feature of digital biomedical images is presense of repeated elements with fuzzy geometry. Cell shape is approximately equals to shape of circle or ellipse. It demands an improved algorithms for region matching. The symmetric difference of regions point sets used to determine the best affine transformation between the regions. Application of region discrete central moments increases the accuracy of contours corresponding points during calculation of affine transformations between the regions of microobjects.

Key words: biomedical image, structural analysis, region matching, central moments, contour, correspondence points.

Вступ

Сьогодні опрацювання зображень є важливим напрямком застосування сучасної обчислювальної техніки. Задачами опрацювання зображень є опис, аналіз та оброблення зображень. Проблеми аналізу зображень крім класичної задачі розпізнавання фігур заданої форми зумовлюють нові завдання опису структури зображення.

Біомедичні зображення – це растрові зображення, отримані за допомогою будь-якої біомедичної техніки, що використовуються для візуального та автоматизованого аналізу у медицині та біології. Для діагностики одного з найпоширеніших захворювань – раку, використовується світлова мікроскопія. Цей метод дослідження є джерелом цифрових цитологічних і гістологічних зображень. Такі зображення містять від десятків до сотень клітин одного або декількох типів, що дозволяє, в контексті комп'ютерного зору, розглядати їх як зображення повторюваних елементів. Зображення повторюваних елементів є предметом аналізу в системах штучного інтелекту та широко використовуються в комп'ютерній графіці. Ці зображення можна описати так:

- 1) існує фрагмент який регулярно повторюється в межах області, яка є великою порівняно з його розміром;
- 2) цей фрагмент утворено елементарними (непохідними) частинами, що розміщені певним способом;
- 3) елементарні частини є приблизно однаковими в межах всього зображення.

Аналіз зображень повторюваних елементів застосовується [1, 2] під час опрацювання дистанційно отриманих зображень (аерокосмічних знімків), контролю якості матеріалів і тканин, оброблення документів та ін.

Нижче розглянуто основні результати в області аналізу цифрових гістологічних зображень (ЦГЗ). Вибір методів аналізу ЦГЗ зумовлений, насамперед, варіабельністю ознак мікрооб'єктів – площі, периметра, округлості тощо.

В області сегментації кольорових цитологічних зображень автори [3] розробили алгоритм розділення зображення на кольорові складові, що дозволило виділити ядро та цитоплазму. Алгоритм полягає в обчисленні внеску кожного із застосовуваних фарбників, базуючись на їх спектрі поглинання світла. Для кожного типу фарбника наперед обчислено вектор кольорів у RGB просторі.

В області класифікації типу клітин задачу постановки діагнозу, а саме висновку про наявність злякисного новоутворення, вирішено для окремих видів пухлин [4, 5]. У роботі [4] детально показано кроки постановки діагнозу. Інтерпретація у виконанні спеціаліста патологоанатома – повинна спричинити збирання морфологічних доказів (числових характеристик об'єктів), що дозволяють ідентифікувати патологічні процеси, які можуть виявитися доброякісними або злякисними. Цей крок передбачає досягнення трьох різних завдань: морфологічного аналізу, ідентифікації клітин, постановки та інтерпретації діагнозу.

Питання побудови нечітких правил для постановки діагнозу на основі якісних ознак описано в роботі [4]. Умовами правил виступають ознаки окремих мікрооб'єктів (клітин).

Із наведеного вище видно, що задачі діагностування розв'язані для частини патологічних процесів, що розглядають як мікрооб'єкт клітину. Проте існує багато задач діагностування патологічних процесів, що базуються на аналізі характеристик шарів тканин та інших складних об'єктів. Тому актуальною є задача групування клітин у складні об'єкти на основі їх взаємного просторового розміщення, тобто структурних залежностей.

1. Зіставлення областей під час структурного аналізу зображень

Відомо, що проблема відповідності (кореспонденції) форми областей є однією з найбільш складних проблем в обробці зображень не тільки через ітеративний пошук, але й чутливість до шуму [6]. В загальному випадку відповідність форми областей I та J , представленими множинами точок, задається відображенням $\forall i \in I, \exists j \in J: T(i) = j$ [7].

При пошуку повторюваних елементів використовують [8] кореляцію знайденого елемента з зображенням для пошуку параметрів перетворення. У роботі [9] для зіставлення повторюваних елементів застосовується кореляція попередньо перетвореного елемента з елементом кандидатом. Такий підхід успішно використовується для зображень об'єктів штучної природи. Проте мікрооб'єкти одного класу (клітини, ворсинки) на цифрових гістологічних зображеннях відрізняються за формою, кольором, текстурою і розміром у певних межах. Тому кореляційний метод зіставлення та метод на основі характерних точок контуру дають неточний результат.

Для опису структури зображення використовуються такі зв'язки між сегментами: сусідство, трансляція, дотик, включення, перекриття, просторові відношення. Однією з найбільш часто застосовуваних структур є граф сусідства областей, кожна вершина якого представляє певну область зображення. Дві вершини зв'язані ребром тоді і лише тоді, коли відповідні області мають спільну границю. Граф сусідства створюється на основі розмітки шляхом прослідковування границі кожної області та зберігання ідентифікаційних міток сусідніх областей або у вигляді квадродерева.

У статті [10] запропоновано 8 елементарних топологічних відношень (рис. 1, а): роз'єднані, ззовні пов'язані, рівні, частково перекриваються, дотична власна частина, інверсна дотична власна частина, недотична власна частина, інверсна недотична власна частина.

Система прямокутних відношень головних напрямків використовує відношення між обмежуючими прямокутниками областей [11]. Положення областей одна відносно одної визначається напрямками сторін світу, наприклад "північний схід" (рис. 1, б).

Недоліками простих систем просторових відношень є відсутність врахування орієнтації області та її форми.

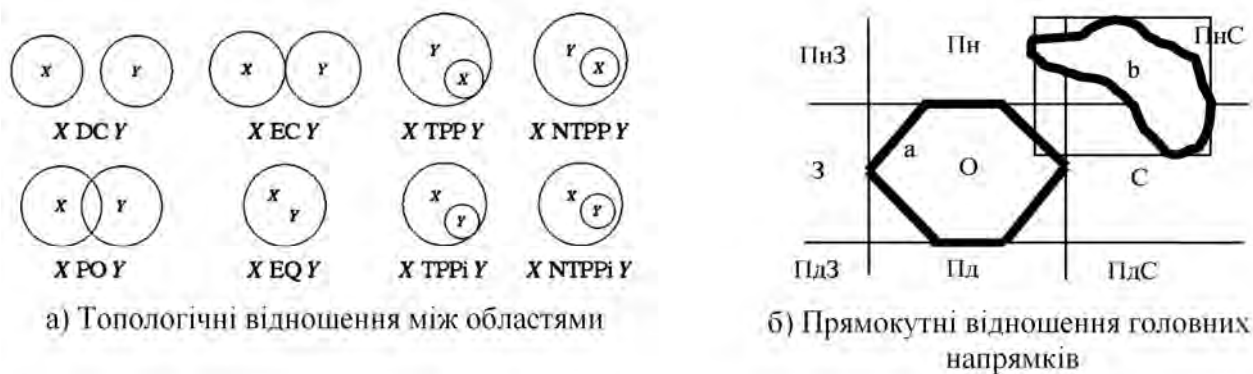


Рис. 1. Системи просторових відношень

Для усунення недоліків застосуємо підхід [12] до опису структурних залежностей на основі операції афінного перетворення T , а саме комбінації трансляції, повороту та масштабування (1)

$$[x' \ y' \ 1] = \begin{bmatrix} a & b & p \\ c & d & q \\ m & n & 1 \end{bmatrix} [x \ y \ 1] = [T][x \ y \ 1], \quad (1)$$

де x, y – координати точки, x', y' – координати перетвореної точки, a, b, c, d – коефіцієнти зсуву і масштабування; m, n – коефіцієнти трансляції, p, q – коефіцієнти проєкцій.

Для знаходження перетворення g між двома зображеннями або областями достатньо взяти три точки, які не лежать на одній прямій, на області A і знайти відповідні їм точки області B . Ця вимога справедлива для областей *однакової* форми (або для двох однакових зображень). В цьому випадку області пов'язані тільки поворотом і трансляцією.

Характерною властивістю цифрових біомедичних зображень є зміна форми об'єктів одного класу (клітин) та наближення їх форми до форми еліпса або до форми фігури з елементом n -кратної симетрії повороту. Тобто пряма відповідність між точками двох областей при афінному перетворенні неможлива, навіть враховуючи похибку дискретизації.

Для оцінки відмінності областей одна від одної використовується розмір області перекриття [13]. Для двох компактних множин A і B область перекриття M визначено як симетричну різницю

$$M = \Delta B = (A \setminus B) \cup (B \setminus A), \quad (2)$$

де Δ – операція симетричної різниці.

Ця міра відмінності є метрикою. У випадку плоских областей одного класу з нечіткою геометричною структурою прийемо, що афінне перетворення g^* переводить область A у B так, щоб мінімізувати симетричну різницю відповідних множин точок:

$$g^* = \arg \min_g (\Delta \Delta(g^{-1}B)). \quad (3)$$

Для знаходження трьох відповідних точок можна застосовувати простий алгоритм на основі серединного перпендикуляра [14]. Він полягає в знаходженні координат відрізка максимальної довжини. Третя точка знаходиться в місці перетину контуру і серединного перпендикуляра побудованого на відрізку. Прикладом комбінаторного методу розв'язання задачі пошуку найкращого перетворення є метод січних прямих [6]. Він полягає у побудові прямих, які перетинають контур, з деякої стартової точки через заданий кут α . Із множини точок перетину знаходяться три точки, які мінімізують похибку перетворення між відповідними координатами на заданих контурах.

Вимогами до *відповідних* точок є стабільність їх положення на контурі – інваріантність вибору по відношенню до повороту, трансляції, спотворень контуру тощо. Відповідні точки обираються з множини характерних точок контуру (ХТ), які відображають кривизну контуру, точки максимуму і мінімуму і т. д. Слабко виражені кутові ознаки контурів мікрооб'єктів та відмінність форми клітин вимагає додаткового застосування алгоритмів визначення найкращої відповідності

між самими ХТ [7]. Повторюваність ХТ на контурі, за умов дії на область афінних перетворень, обчислюється як відношення кількості спотворених зображень із однаково визначеними ХТ до всієї кількості тестових зображень. Для дослідження впливу ознак форми мікрооб'єктів на повторюваність ХТ використаємо ознаку округлості. Округлість області F_O розраховується так:

$$F_O = \frac{4\pi S}{P^2},$$

де S – площа області, P – периметр.

Проведено тестування алгоритму визначення точок максимальної кривизни [15] на областях різних класів. Тестовими взято зображення об'єктів із аерокосмічних зображень (АЗ), зображень документів (ЗД) та цитологічних зображень (ЦЗ1, ЦЗ2), які показані на рис. 2.

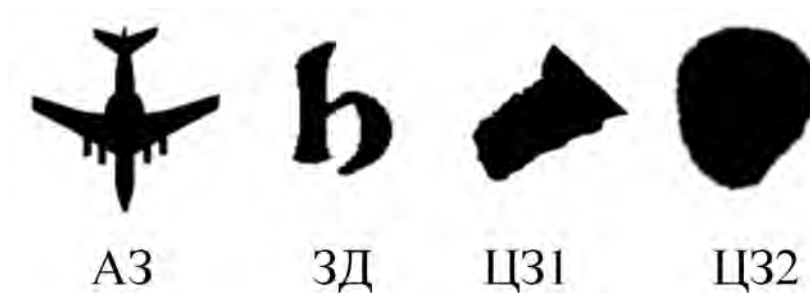


Рис. 2. Тестові зображення

Тестові зображення відрізняються кривизною контуру та ознакою округлості (табл. 1).

Таблиця 1

Повторюваність визначення ХТ

Клас зображення	АЗ	ЗД	ЦЗ1	ЦЗ2
Округлість	0,113	0,1705	0,51	0,8375
Рівень повторюваності ХТ	0,94	0,88	0,91	0,92

Відсутність кутових ознак приводить до неправильного визначення відповідних точок навіть на одному і тому ж контурі при дії повороту і трансляції на область. Для розв'язання задачі пошуку афінного перетворення можливим рішенням є обчислення моментів області, які характеризують область, проте обчислюються не на основі ознак контуру.

Метою статті є удосконалення методу визначення відповідних точок на контурах мікрооб'єктів біомедичної природи.

2. Визначення відповідних точок на основі центральних моментів

Розглянемо мікрооб'єкти на бінарному зображенні, заданому функцією

$$b(x, y) = \begin{cases} 1 & \text{об'єкт} \\ 0 & \text{фон} \end{cases}.$$

Наведемо визначення просторових моментів [16] для бінарних зображень:

$$m_{p,q} = \iint x^p y^q b(x, y) dx dy, \quad (4)$$

де p, q – індекси порядку моменту.

Моменти класифікують за їх порядком, тобто сумою $p + q$. Момент нульового порядку $((p, q) = (0,0))$ описує площу об'єкта. Моменти першого порядку $((p, q) = (1,0)$ або $(0,1))$ містять інформацію про центр мас об'єкта:

$$x_c = \frac{m_{1,0}}{m_{0,0}}, \quad y_c = \frac{m_{0,1}}{m_{0,0}}.$$

Визначення (4) описує загальні просторові моменти об'єкта. Із просторових моментів центральні моменти можуть бути отримані скороченням за допомогою центра мас (x_c, y_c) області. Центральні моменти обчислюються так:

$$\mu_{p,q} = \iint (x - x_c)^p (y - y_c)^q b(x, y) dx dy$$

Центральні моменти першого або вищого порядку можуть безпосередньо бути отримані із просторових моментів:

$$\mu_{p,q} = \frac{m_{p,q}}{m_{0,0}} - \left(\frac{m_{1,0}}{m_{0,0}} \right)^p * \left(\frac{m_{0,1}}{m_{0,0}} \right)^q$$

Центральні моменти другого порядку можна виразити так:

$$\mu_{2,0} = \frac{m_{2,0}}{m_{0,0}} - \left(\frac{m_{1,0}}{m_{0,0}} \right)^2 = \frac{m_{2,0}}{m_{0,0}} - x_c^2,$$

$$\mu_{0,2} = \frac{m_{0,2}}{m_{0,0}} - \left(\frac{m_{0,1}}{m_{0,0}} \right)^2 = \frac{m_{0,2}}{m_{0,0}} - y_c^2.$$

На основі центральних моментів другого порядку обчислюється [16] значення головних осей інерції:

$$\lambda_{1,2} = \sqrt{\frac{1}{2} * (\mu_{2,0} + \mu_{0,2}) \pm \sqrt{4 * \mu_{1,1}^2 - (\mu_{2,0} - \mu_{0,2})^2}}.$$

Головні осі інерції об'єкта відповідають великій і малій осям a і b еліпса зображення, який може бути використаний як апроксимація об'єкта. Головні осі інерції – це ті осі, навколо яких об'єкт може обертатися з мінімальною (велика вісь a) або максимальною (мала вісь b) інерцією. Осі a і b показані на рис. 3.

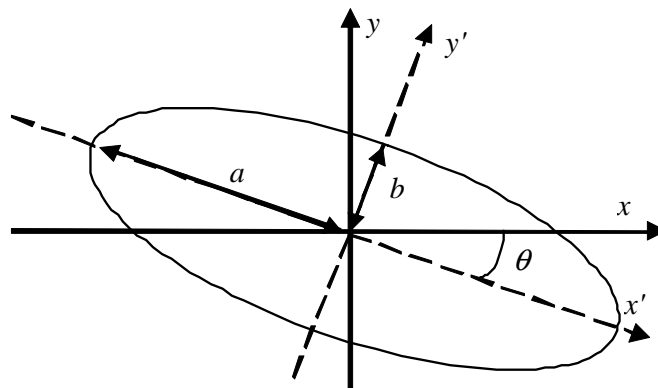


Рис. 3. Еліпс області

Орієнтація об'єкта визначається як кут нахилу осі, довкола якої об'єкт може обертатися з мінімальною інерцією, до осі абсцис (тобто кут нахилу великої осі). У цьому напрямі область має своє найбільше видовження. Орієнтація області обчислюється так:

$$\theta = \frac{1}{2} \operatorname{arctg} \frac{2\mu_{1,1}}{\mu_{2,0} - \mu_{0,2}}.$$

Існує двозначність у визначенні кута нахилу θ положення об'єкта (рис. 3), яку вирішено шляхом вибору кута між віссю абсцис і великою віссю a . Кут нахилу θ набуває значень відповідно до табл. 2.

На основі обчислених ознак потрібно визначити три ХТ на контурі об'єкта. Ці точки повинні визначатися за правилами, що дозволяють знаходити відповідні їм точки на інших об'єктах.

Кут нахилу еліпса зображення

$\mu_{2,0} - \mu_{0,2}$	$\mu_{1,1}$	θ	
Нуль	Нуль	0°	
Нуль	Додатне	$+45^\circ$	
Нуль	Від'ємне	-45°	
Додатна	Нуль	0°	
Від'ємна	Нуль	-90°	
Додатна	Додатне	$\frac{2}{2} \arctan \frac{2\mu_{1,1}}{\mu_{2,0} - \mu_{0,2}}$	$0^\circ < \theta < 45^\circ$
Додатна	Від'ємне	$\frac{1}{2} \arctan \frac{2\mu_{1,1}}{\mu_{2,0} - \mu_{0,2}}$	$-45^\circ < \theta < 0^\circ$
Від'ємна	Додатне	$\frac{1}{2} \arctan \frac{2\mu_{1,1}}{\mu_{2,0} - \mu_{0,2}} + 90^\circ$	$45^\circ < \theta < 90^\circ$
Від'ємна	Від'ємне	$\frac{1}{2} \arctan \frac{2\mu_{1,1}}{\mu_{2,0} - \mu_{0,2}} - 90^\circ$	$-90^\circ < \theta < -45^\circ$

Метод визначення координат ключових точок на контурі складається з таких кроків (рис. 4).

1. Для об'єкта обчислюємо кут нахилу головної осі θ , координати центра мас об'єкта $M(x_c, y_c)$.
2. Будуємо пряму m , яка проходить через центр мас об'єкта і перетинає контур об'єкта. Точки $A(x_1, y_1)$ і $B(x_2, y_2)$ визначено так, щоб виконувалась умова $|AM| > |MB|$.
3. Для визначення третьої точки C через центр мас проводимо пряму n перпендикулярно до великої осі. У випадку простого опуклого контуру пряма n перетинає його в двох точках C і C' . Третю ключову точку C обираємо так, щоб кут між відрізком MB і MC дорівнював 90° .

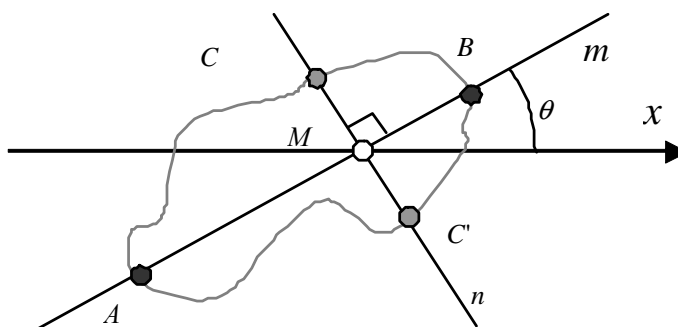


Рис. 4. Відповідні точки на контурі області

Отже, для об'єкта ідентифіковані три ХТ його контуру, на основі яких буде визначатися відповідність між об'єктами.

Схему алгоритму знаходження відповідних точок показано на рис. 5.

Експериментальні дослідження

Загальний алгоритм аналізу ЦГЗ сформовано з таких кроків:

- 1) формуємо тестові зображення із двох мікрооб'єктів шляхом афінного перетворення;
- 2) кольорове зображення перетворюємо у градації сірого;
- 3) проводимо розбиття зображення на фон та області за допомогою алгоритму порогової сегментації;
- 4) виділяємо контури областей, обчислюємо координати центра мас, велику і малу вісь, кут орієнтації області;
- 5) обчислюємо координати трьох відповідних точок на контурі розробленим алгоритмом, алгоритмом серединного перпендикуляру та методом січних;
- 6) обчислюємо коефіцієнти афінного перетворення однієї області в іншу на основі координат відповідних точок;
- 7) проводимо зіставлення областей на основі обчисленого афінного перетворення та обчислення симетричної різниці.

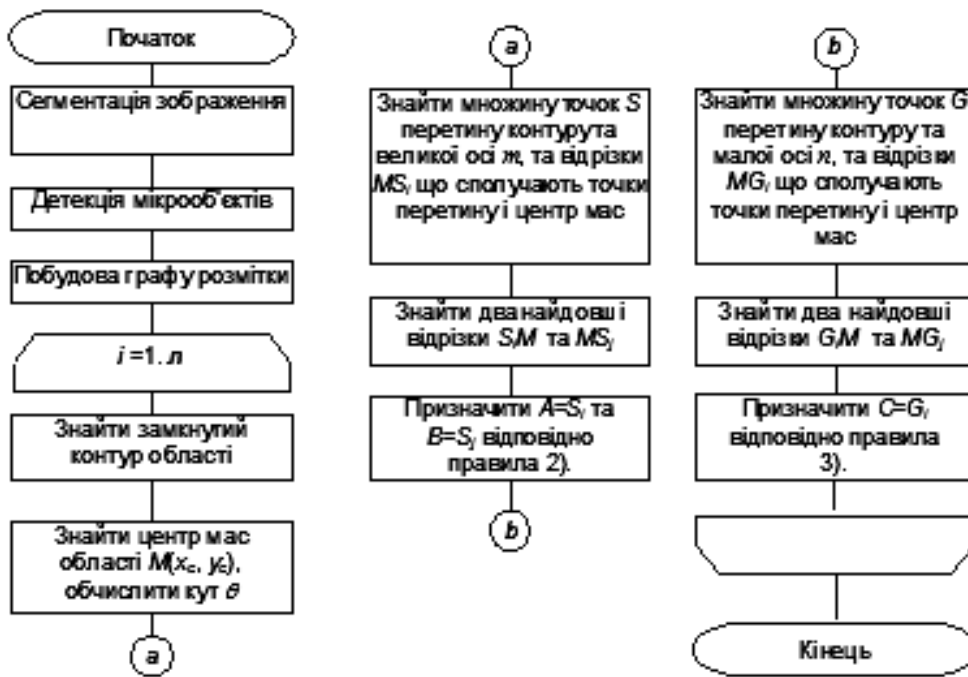


Рис. 5. Схема алгоритму знаходження відповідних точок

Параметри зображення, що впливають на результат виявлення характерних точок такі: поворот зображення, розмиття, зміна масштабу, зменшення освітленості, збільшення коефіцієнта JPEG стиснення, збільшення піксельного шуму. За цими параметрами оцінюють алгоритми, що визначають ХТ на цілому зображенні. Цифрові гістологічні зображення обов'язково підлягають операції попереднього оброблення з вирівнювання нерівномірності фону, спричиненої нерівномірним освітленням. Тому ми не оцінюватимемо ефективність розробленого алгоритму за умови зменшення освітленості.

Для дослідження ефективності розробленого методу проведено його тестування на базі зображень окремих клітин взятих з БД цитологічних і гістологічних зображень Breast Cancer Dataset. Зіставлено області об'єктів одного класу між собою, обчислення найкращого афінного перетворення та симетричної різниці (2). Площу перетину нормовано відносно площі одного із об'єктів. Результати тестування під час дії різних комбінацій перетворень наведені на рис. 6.

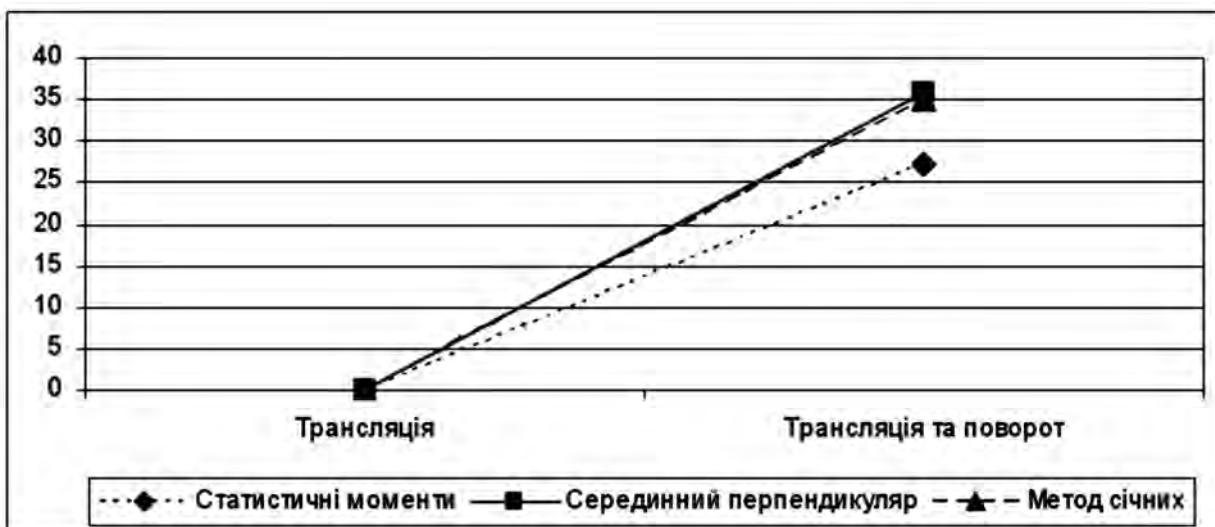


Рис. 6. Величина симетричної різниці для різних алгоритмів

Метод січних з кроком збільшення кута побудови на 10° потребує на порядок більше часу на перебір всіх можливих матриць перетворень ніж розроблений метод. Розроблений метод визначення відповідних точок на основі центральних моментів виявляється кращим за дії цілочисельної трансляції та повороту.

Висновки

Результати експериментів підтверджують доцільність застосування методу визначення відповідних точок на основі дискретних центральних моментів під час розв'язання задачі відповідності двох областей одного класу з різною формою. Основна перевага центральних моментів, а саме їх інваріантність до трансляції і повороту області, дозволила зменшити величину перекриття зіставлених областей за допомогою розробленого методу на 8 % порівняно із методом січних.

1. *Handbook of Texture Analysis* / Ed. by Majid Mirmehdi, Xianghua Xie, Jasjit Suri. – London, UK : Imperial College Press, 2008. – 413. P. 2. Tuceryan M. *Handbook of Pattern Recognition and Computer Vision* / M. Tuceryan, A.K. Jain // *Handbook of Pattern Recognition and Computer Vision (2 nd Edition)* / Ed. by C. H. Chen, L.F. Pau, P.S.P. Wang. – World Scientific Publishing Co., 1998. – P. 207–249. 3. Ruifrok A. C. *Quantification of histochemical staining by color deconvolution* / A. C. Ruifrok, D. A. Johnston // *Analytical and quantitative cytology and histology. The International Academy of Cytology and American Society of Cytology*. – 2001. – Vol. 23, no. 4. – P. 291–299. 4. Ovalle A. *Kids: A distributed expert system for biomedical image interpretation.* / A. Ovalle, Catherine Garbay // *Information Processing in Medical Imaging* / Ed. by Alan C. F. Colchester, David J. Hawkes. – Vol. 511 – 1991. – P. 419–433. 5. *Characterization of chromatin texture by contour complexity for cancer cell classification* / T. Kiyuna, A. Saito, E. Kerr, W. Bickmore // *8th IEEE International Conference on Bioinformatics and BioEngineering, 2008.* – Athens, 2008. – P. 1–6. 6. Путьятин Е.П. *Обработка изображений в робототехнике* / Е. П. Путьятин, С. И. Аверин. – М.: Машиностроение, 1990. – 320 С. 7. *Contour correspondence via ant colony optimization* / O. van Kaick, G. Hamarneh, H. Zhang, P. Wighton // *Proc. 15th Pacific Conference on Computer Graphics and Applications (Pacific Graphics 2007)* – Maui, Hawaii, United States, 2007. – P. 271–280. 8. Schaffalitzky F. *Geometric grouping of repeated elements within images* / F. Schaffalitzky, A. Zisserman // *In Proc. 9th British Machine Vision Conference (BMVC), Southampton, September 1998.* – Southampton, UK: Springer-Verlag, 1998. – September. – P. 13–22. 9. Liu Y. *A computational model for periodic pattern perception based on frieze and wallpaper groups* / Yanxi Liu, Robert T. Collins, Yanghai Tsing // *IEEE Transactions on Pattern Analysis and Machine Intelligence (TPAMI)*. – 2004. – Vol. 26, No. 1. – P. 354–371. 10. Wolter F. *Spatio-temporal representation and reasoning based on rcc-8* / Frank Wolter, Michael Zakharyashev // *In Proceedings of the seventh Conference on Principles of Knowledge Representation and Reasoning (KR2000)*. – Breckenridge, Colorado, USA: Morgan Kaufmann, 2000. – P. 3–14. 11. Navarrete I. *Consistency checking of basic cardinal constraints over connected regions* / Isabel Navarrete, Antonio Morales, Guido Sciavicco // *IJCAI 2007, Proceedings of the 20th International Joint Conference on Artificial Intelligence, Hyderabad, India, January 6-12, 2007.* – 2007. – P. 495–500. 12. Березский О.Н. *Теоретико-групповой подход к анализу симметричных изображений* / О. Н. Березский // *Управляющие машины и системы.* – 2010. – Т. 2. – С. 16–24. 13. Veltkamp R. C. *State of the art in shape matching* / Remco C. Veltkamp, Michiel Hagedoorn // *Principles of visual information retrieval.* – 2001. – P. 87–119. 14. Березский О.Н. *Алгоритмы анализа и синтеза биомедицинских изображений* / О. Н. Березский // *Проблемы информатики и управления.* – 2007. – № 2. – P. 134–144. 15. Chetverikov D. *A Simple and Efficient Algorithm for Detection of High Curvature Points in Planar Curves* / Dmitry Chetverikov / Ed. by *Computer Analysis of Images, Patterns.* – 2004. – Vol. 2756/2003 – P. 746–753. 16. Kilian J. *Simple image analysis by moments. OpenCV library documentation [Електронний ресурс].* / J. Kilian – 2001. – Режим доступу до статті.: <http://public.cranfield.ac.uk/~c5354/teaching/dip/opencv/SimpleImageAnalysisbyMoments.pdf>.