

УДК 004.08

В.В. ГРИЦИК

Тернопільський національний технічний університет імені Івана Пулюя
Національний університет "Львівська політехніка"

А.Я. ДУНАС

Національний університет "Львівська політехніка"

ДОСЛІДЖЕННЯ МЕТОДІВ РОЗПІЗНАВАННЯ ОБРАЗІВ ДЛЯ СИСТЕМ КОМП'ЮТЕРНОГО ЗОРУ РОБОТІВ МАЙБУТНЬОГО

Представлено дослідження елемента машино-машинного інтерфейсу на предмет можливості динамічної адаптації для покращення сприйняття зовнішнього середовища шляхом розробки методики адаптації робозору до візуального спектру. Досліджено і порівняно між собою різні порогові методи сегментації зображень різних категорій.

Ключові слова: сегментація зображень, робозір, бінаризація, розпізнавання образів, метод Ніблака, порогова бінаризація, гістограмні методи, метод Оцу, метод Єна, метод трикутника, роботи майбутнього.

В.В. ГРИЦИК

Тернопольский национальный технический университет имени Ивана Пулюя
Національний університет "Львівська політехніка"

А.Я. ДУНАС

Национальный университет "Львовская политехника"

ИССЛЕДОВАНИЕ МЕТОДОВ РАСПОЗНАВАНИЯ ОБРАЗОВ ДЛЯ СИСТЕМ КОМПЬЮТЕРНОГО ЗРЕНИЯ РОБОТОВ БУДУЩЕГО

Представлено исследование элемента машино-машинного интерфейса на предмет возможности динамической адаптации для улучшения восприятия внешней среды путем разработки методики адаптации роботизированного зрения к визуальному спектру. Исследованы и сравнены между собой различные пороговые методы сегментации изображения различных категорий.

Ключевые слова: сегментация изображений, роботизированное зрение, бинаризация, распознавания образов, метод Ниблака, пороговая бинаризация, гистограммные методы, метод Оцу, метод Єна, метод треугольника, роботы будущего.

V.V. HRYTSYK

Ternopil Ivan Puluji National Technical University
Lviv Polytechnic National University

A.Y. DUNAS

Lviv Polytechnic National University

RESEARCH OF PATTERN RECOGNITION METHODS FOR COMPUTER VISION SYSTEMS OF ROBOTS OF THE FUTURE

The research of the machine-machine interface element concerning the possibility of dynamic adaptation for improving the perception of the visual environment by developing a method for adapting the computer vision to the visual spectrum. The various threshold methods of image segmentation from different categories are investigated and compared among themselves.

Keywords: image segmentation, computer vision, binarization, pattern recognition, Niblak method, threshold binarization, histogram methods, Otsu method, Yen's method, Triangle method, robots of tomorrow.

Постановка проблеми

Комп'ютерні технології займають настільки важливу роль в житті сучасного суспільства, що зараз вже важко уявити собі жоден вид діяльності, в якому так чи інакше не використовуються комп'ютери [3]. Ми живемо в період 4-ої промислової революції, і прогрес знань і технологій різко пришвидшується з кожним днем, адже створення одних новітніх технологій виступає фундаментом для появи нових. Вже зараз у численних галузях промисловості, науки, інтернет-технологій тощо ефективно використовуються розробки, пов'язані з використанням систем штучного інтелекту, а також різного виду роботів чи роботизованих систем.

Роботи майбутнього – це технології з інтелектуальними візуальними можливостями: вони повинні відчувати і розуміти динаміку реального світу [7]. Тому розробка систем комп'ютерного зору з динамічним

сприйняттям зовнішнього візуального середовища є критично важливою задачею для роботів наступної генерації. А це потребуватиме розробки відповідних моделей і засобів синтезу методів сприйняття даних візуального спектру, що надходять в режимі реального часу [2, 4]. Отже, актуальною проблемою розвитку штучного інтелекту є розробка принципів сприйняття комп'ютером зовнішнього світу через розуміння відео даних [1, 5].

Формулювання мети дослідження

Метою роботи є покращення сприйняття візуального світу робозором шляхом розробки методики його адаптації до особливостей зовнішнього середовища.

Викладення основного матеріалу дослідження

Дослідимо та порівняємо декілька найбільш поширених методів сегментації візуального поля уваги, що працюють у реальному часі та ґрунтуються на різних математичних апаратах обробки інформації.

1. Бінаризація з нижнім порогом.

Цей метод відноситься до групи методів сегментації зображення, що ґрунтуються на глобальних порогах. Суть методу полягає в тому, що глобальний поріг яскравості вибирається як деяка константа, і в залежності від відношення цієї константи до локальних значень яскравості для кожного пікселя, бінаризується значення цього пікселя. Цей метод є найбільш простим і поширеним. Опишемо його наступним чином:

$$a_{xy}^{new} = \begin{cases} 0, & \text{if } \frac{r+g+b}{3}(a_{xy}) \leq L \\ 1, & \text{if } \frac{r+g+b}{3}(a_{xy}) > L \end{cases} \quad (1.1)$$

де a_{xy}^{new} – результуюче значення яскравості пікселя;
 a_{xy} – вхідне значення яскравості пікселя;
 $L(const) \in [0, 255]$ – глобальний рівень освітлення;
 r, g, b – оригінальні значення червоної, зеленої та синьої складової кольору пікселя a_{xy} .
 Візуалізацію див. на рис. 6.1.б).

2. Метод Ніблака

Метод ґрунтується на обчисленні локального порогу освітленості.

Ідея полягає у вирівнюванні порогу яскравості бінаризації від точки до точки на основі відхилення **локального середнього** значення яскравості (значення, що обчислюється для кожного пікселя, базуючись на значеннях яскравості його самого і його сусідів), від **локального** (обчислюється лише для одного пікселя) в заданій масці [6].

Тобто, бінаризоване представлення пікселя обчислюється так:

$$a_{xy}^{new} = \begin{cases} 0, & \text{if } B(x, y) \leq L \\ 1, & \text{if } B(x, y) > L \end{cases} \quad (2.1)$$

де $B(x, y) \in [0, 255] = \frac{r+g+b}{3}(a_{xy})$, – локальне значення яскравості пікселя a_{xy} ;
 $L \in [0, 255] = m_{w \times w}(x, y) + k * s_{w \times w}(x, y)$ – локальний поріг яскравості для пікселя a_{xy} в околі $w \times w$;
 $m_{w \times w}(x, y) \in [0, 255] = \frac{\sum_1^{w \times w} B(x, y)}{w \times w}$ – середнє значення яскравості в околі пікселя $w \times w$;
 $s_{w \times w}(x, y) = \sqrt{\frac{1}{w \times w} \sum_1^{w \times w} (B(x, y) - m_{w \times w}(x, y))^2}$ – середньоквадратичне відхилення вибірки в околі пікселя;
 $k(const) = -0,2$ для об'єктів, що імовірно представлені чорним кольором (тобто, якщо $B(x, y) \leq 127$), і $k = 0,2$ для об'єктів, що імовірно представлені білим кольором $B(x, y) > 127$;
 $w(const)$ – розмір маски околу, наприклад, 15 [6].
 Візуалізацію див. на рис. 6.1.в).

3. Метод Оцу

Метод використовує гістограму розподілу значень яскравості пікселів зображення.

Суть метода Оцу полягає в тому, щоб виставити поріг між класами на гістограмі таким чином, аби кожен з них був якомога "щільнішим". Якщо висловлюватись математичною мовою, то це зводиться до мінімізації внутрішньокласової дисперсії, яка визначається як зважена сума дисперсій двох класів:

$$\sigma_w^2(L) = w_1(L)\sigma_1^2(L) + w_2(L)\sigma_2^2(L), \quad (3.1)$$

де ваги w_i – це ймовірності двох класів, що розділені порогом L ;
 σ_i^2 – дисперсії цих класів.

Оцу довів, що мінімізація дисперсії *всередині* класу рівносильна максимізації дисперсії *між* класами, яку можна виразити через ймовірність w_i і середній арифметичний клас μ_i :

$$\sigma_b^2(L) = \sigma^2 - \sigma_w^2(L) = w_1(L)w_2(L)[\mu_1(L) - \mu_2(L)]^2. \quad (3.2)$$

Для початку потрібно побудувати гістограму $p(l)$ зображення і визначити частоту входжень $N(l)$ кожного рівня яскравості зображення $G(x, y)$. Також шукаємо сумарну яскравість N_T пікселів зображення:

$$N_T = \sum_{i=0}^{\max(G)} p(i) \quad (3.3)$$

Після цього для кожного значення напівтону (порогу) $L = \overline{1, \max(G)}$ виконуємо наступне:

$$\omega_1(L) = \frac{\sum_{i=0}^{L-1} p(i)}{N_T} = \sum_{i=0}^{L-1} N(i), \omega_2(L) = 1 - \omega_1(L) \quad (3.4)$$

$$\mu_T = \frac{\sum_{i=0}^{\max(G)} i * p(i)}{N_T} = \sum_{i=0}^{\max(G)} i \times N(i) \quad (3.5)$$

$$\mu_1(L) = \frac{\sum_{i=0}^{L-1} i * p(i)}{N_T * \omega_1(L)} = \frac{\sum_{i=0}^{L-1} i * N(i)}{w_1(L)} \mu_2(L) = \frac{\mu_T - \mu_1(L) * \omega_1(L)}{w_2(L)} \quad (3.6)$$

$$\sigma_b^2(L) = \sigma^2 - \sigma_w^2(L) = w_1(L)w_2(L)[\mu_1(L) - \mu_2(L)]^2, \quad (3.7)$$

Шуканий поріг дорівнює значенню L , при якому $\sigma_b^2(L)$ максимальне:

$$L = \operatorname{argmax}_L \sigma_b^2(L) \quad (3.8)$$

Бінаризоване представлення пікселя обчислюється так:

$$a_{xy}^{new} = \begin{cases} 0, & \text{if } B(x, y) \leq L \\ 1, & \text{if } B(x, y) > L \end{cases} \quad (3.9)$$

Візуалізацію див. на рис. 6.1.г).

4. Метод Єна

Цей метод відноситься до методів, які використовують ентропію розподілу яскравості кольорів на зображенні. Метод Єна розглядає об'єкт на зображенні і фон, на якому цей об'єкт розташований, як два різні джерела зорової інформації. І значення яскравості, при якому сума цих двох ентропій сягає свого максимуму, вважається оптимальним порогом для сегментації зображення [8].

Для початку потрібно обчислити гістограму $p(l)$ зображення і частоту входжень $N(l)$ кожного рівня яскравості зображення $G(x, y)$. Також шукаємо сумарну яскравість N_T пікселів зображення:

$$N_T = \sum_{i=0}^{\max(G)} p(i) \quad (4.1)$$

Будуємо допоміжні нормалізовані гістограми:

$$p_{norm}(i) = \frac{p(i)}{N_T}, \quad (4.2)$$

$$p_{normC}(i) = p_{normC}(i - 1) + p_{norm}(i), \quad (4.3)$$

$$p'_{norm}(i) = p'_{norm}(i - 1) + p_{norm}(i)^2, \quad (4.4)$$

$$p''_{norm}(i) = p''_{norm}(i + 1) + p_{norm}(i + 1)^2, \quad (4.5)$$

Знаходимо ентропії об'єкта і його фону:

$$C_f(T) = -\log\{p_{normC}(i) \times (1 - p_{normC}(i))\} \quad (4.6)$$

$$C_b(T) = -\log\{p'_{norm}(i) \times p''_{norm}(i)\} \tag{4.7}$$

Визначаємо значення i , при якому сума цих ентропій максимальна:

$$L = \operatorname{argmax}_i\{C_b(T) + C_f(T)\} \tag{4.10}$$

Використовуємо це значення як поріг яскравості і бінаризуємо зображення:

$$a_{xy}^{new} = \begin{cases} 0, & \text{if } B(x, y) \leq L \\ 1, & \text{if } B(x, y) > L \end{cases} \tag{4.11}$$

Візуалізацію див. на рис. 6.1.д).

5. Метод трикутника

Метод використовує гістограму розподілу значень яскравості на зображенні.

На гістограмі проводиться лінія s від найменшого b_{min} значення яскравості до найбільшого b_{max} . Порогом визначається елемент гістограми, відстань від якого до s є найбільшою [9]:

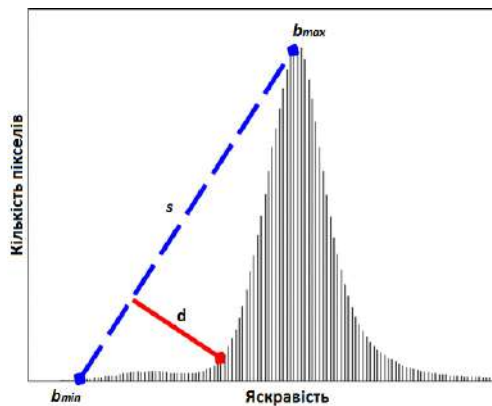


Рис. 5.1. Метод трикутника

$$L = \operatorname{argmax}_{b(i)} d, \text{ де} \tag{5.1}$$

L – значення порогу;

d – відстань від значення гістограми $b(i)$ до s .

Проводимо бінаризацію за стандартною формулою:

$$a_{xy}^{new} = \begin{cases} 0, & \text{if } B(x, y) \leq L \\ 1, & \text{if } B(x, y) > L \end{cases} \tag{5.2}$$

Візуалізацію див. на рис. 6.1.е).

6. Оцінка роботи методів бінаризації

Для оцінки правильності роботи наведених вище методів було проведено кореляцію результатів методом MSE. Як еталонне зображення було використано оригінальне зображення, бінаризоване методом з нижнім порогом бінаризації. Досліджуваним виступало зображення, отримане внаслідок обробки оригінального кожним з методів. Похибка обчислювалась за формулою:

$$MSE = \frac{1}{w \cdot h} \sum_{i=0}^{w-1} \sum_{j=0}^{h-1} |I(i, j) - K(i, j)|^2 \tag{6.1}$$

де $I(i, j)$ – значення яскравості пікселя у точці (i, j) досліджуваного зображення;

$K(i, j)$ – значення яскравості пікселя у точці (i, j) еталонного зображення;

w – ширина зображень;

h – висота зображень;

MSE – значення похибки.

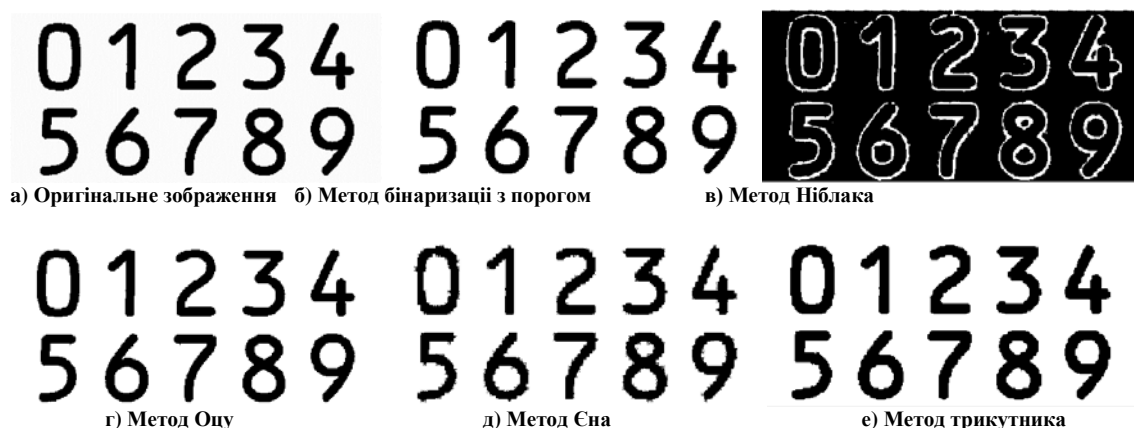


Рис. 6.1. Візуальне порівняння роботи алгоритмів

Висновки

В результаті проведених експериментів були отримані такі результати:

Таблиця 1

Порівняння результатів

№ експерименту	1	2	3	4	5	6	7	8	9	Похибка MSE
Бінаризація з нижнім порогом	3088	3566	3606	3062	3298	3125	3144	3251	3390	±3281
Метод Ніблака	1004	604	874	599	780	905	588	820	991	±796
Метод Оцу	3092	3299	3593	3152	3237	3508	3109	3023	3495	±3279
Метод Єна	780	890	595	773	530	871	642	903	824	±756
Метод трикутника	634	1012	650	931	742	892	638	659	712	±763

Можна зробити висновок, що найкраще у нашому дослідженні себе проявили гістограмні методи бінаризації. Проте для всіх методів дуже велику роль відіграє наближеність вхідного зображення до еталонного, а також наявність побічних "шумів" на ньому. Щодо першої проблеми, вона може бути розв'язана шляхом більш довгого та багатоваріативного навчання системи. Друга проблема відносно легко вирішується шляхом накладення на зображення додаткових фільтрів.

Список використаної літератури

1. Грицик В. В. Оцінка якості передавання і комп'ютерна обробка даних образів / Грицик В. В. // Доповіді НАН України. – 2008. – № 9. Інформатика та кібернетика. – С. 43-48.
2. Audio-visual answer to modern computing // Research*eu results supplement. – 2010. – №26. – P. 31-32p.
3. Мічо Кайку. Фізика майбутнього:Переклала з англ. Анжела Кам'янець. – Львів: Літопис, 2013. – 432 с.
4. Software: running commentary for smarter surveillance? // Reasearch*eu results supplement. – 2010. – №24. – P. 29
5. Hrytsyk V. Augmented reality for people with disabilities / V. Hrytsyk, A. Grondzal, A. Bilenkij // Proceeding soft of the Xth International Scientific and Technical Conference "Computer Sciences and Information Technologies (CSIT'2015) " (Lviv, 14-17 September 2015). – Lviv: IEEE, 2015. – P. 188-191.
6. Korzynska A. Validation of various adaptive threshold methods of segmentation applied to follicular lymphoma digital images stained with 3,3'-Diaminobenzidine&Haematoxylin [Electronic resource] / A. Korzynska, L. Roszkowiak, C. Lopez, R. Bosch and other // Diagnostic Pathology – 2013. – Vol. 8. – I. 48. – Access Mode: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3656801/>
7. [Electronic resource] – Vol. – P. – Access Mode:
8. Sezgin Mehmet. Survey over image thresholding techniques and quantitative performance evaluation / Mehmet Sezgin, Bulent Sankur // Journal of Electronic Imaging. – 2004. – Vol. 13. – I. 1. – P. 146-165.
9. Zack G.W. Automatic Measurement of Sister Chromatid Exchange Frequency / G.W. Zack, W.E. Rogers, S.A. Latt // Journal of Histochemistry and Citochemistry. – 1977. – Vol. 25 – № 7. – P. 741-753.