

А. Клювак, Д. Пелешко, О. Маковейчук, І. Ізонін
Національний університет "Львівська політехніка",
кафедра інформаційних технологій видавничої справи

ДВОЕТАПНЕ СЕГМЕНТУВАННЯ ЗОБРАЖЕННЯ ІЗ СКЛАДНИМ ФОНОМ НА ОСНОВІ МЕТОДУ ОТСУ

© Клювак А., Пелешко Д., Маковейчук О., Ізонін І., 2016

Розглянуто сегментацію зображення із неоднорідним фоном у контексті загальної задачі розпізнавання тексту. Запропонований метод є двоетапним і ґрунтується на сумісному використанні операцій конвертації в колірному просторі, морфологічній обробці та адаптивній фільтрації за алгоритмом Отсу. Серед підзадач, які розглядаються в межах методу, є такі: видалення складного фону, виділення центрального елемента та його адаптивна фільтрація.

Розроблений метод характеризується високою швидкістю та стійкістю до окремих випадків зашумлення.

Ключові слова: текст, зображення, сегментування, бінаризація Отсу, колірні простори, бінарна маска, морфологічна обробка.

The paper is devoted to the task of segmenting images with complex background as a part of more general task of text recognition. The proposed two staged method is based on colour spaces transforms, morphological processing and Otsu's adaptive filtering. The algorithm involves: complex background detection, central element detection, its adaptive filtering.

Key words: text, image segmentation, Otsu binarization, color spaces, binary mask, morphological processing

Вступ та аналіз літературних джерел

Під час інтелектуальної обробки одиничних зображень часто виникають проблеми сегментування зображення. Задача сегментування у переважній більшості випадків визначається потребою виокремити об'єкт чи об'єкти, які наведені на зображенні. Характерною особливістю цих областей є відповідність виконання наперед визначеному критерію на усіх чи майже усіх пікселях сегментованої ділянки. Наявність критерію та умова його виконання визначають поняття однорідності сегментованої ділянки і зв'язності множини елементів зображення. Наприклад, найяскравішим виразником критеріальної ознаки є так звана порогова сегментація, за якою зіставляються значення функції інтенсивності та задане значення (порог), що дає можливість розбити вхідний клас на групу класів (сегментів).

Сьогодні існують різноманітні класифікації методів сегментації. В загальному випадку розглядають класифікацію методів сегментації, за якою їх поділяють на три групи: статистичні, структурно-ознакові та змішані. У зв'язку із існування величезної кількості методів сегментації у кожній із цих груп набувають популярності класифікації методів за різноманітними характеристиками. Наприклад, за режимами обробки (real time, soft regime, online mode, intelligent analysis), за вхідними даними (окреме зображення, масив зображень чи потік зображень), за спеціалізацією (виділення строго визначених об'єктів) тощо. Не менш популярною є класифікація методів: як таких, що ґрунтуються на машинному навчанні, так і таких, які не використовують його. Принциповою відмінністю в останній класифікації методів є існування різноманітних витрат на процедури навчання при забезпеченні високої якості сегментації на етапі використання.

Основні дослідження в роботі зосереджені в області методів сегментації, які не базуються на машинному навчанні. Існуючі сьогодні методи можна умовно розбити на 5 класів [1]:

- методи, які ґрунтуються на використанні ентропії (алгоритми Pun, Kapur, Li, Shanbak, Yen, Brink, Sahoo, Cheng) [2-8];
- методи, які ґрунтуються на кластеризації (алгоритми Riddler, Lloyd, Kittler, Yani, Yawahar) [9-13];
- атрибутивні методи (алгоритми Tsai, Cheng, Hertz, O’Gorman, Huang, Picaz, Leug) [14-18];
- гістограмні методи (алгоритми Rosenfeld, Sezan, Otsu, Carlotto&Olivo, Ramesh, Gui&Kai) [19-22];
- локально порогові методи (алгоритми Yasuda, White, Niblack, Benrsen, Palumbo, Yanowitz, Kamel, Ob, Sauvola) [23-27];
- кореляційні методи (алгоритми Pal, Abutaleb, Chang, Beghdadi, Friel) [28-30].

У наведеній класифікації наведено основні (без модифікацій) методи сегментації за авторами їх розроблення. Відсутність процедури навчання робить усі ці методи достатньо швидкими в реалізації та роботі. Проте для успішного їх використання необхідно враховувати притаманні їм обмеження і недоліки.

Задача сегментації суттєво ускладнюється у випадку сегментування множини об’єктів, кожен з яких може мати свої параметри яскравості чи зашумлення. Однією із найвідоміших задач сегментування множини об’єктів є виділення тексту, зображеного на деякому фоні. Без цього фону задача сегментації розв’язується процедурою бінаризації.

Суть процедури бінаризації [31] полягає у приведенні існуючого зображення до зображення із бінарними значеннями функції інтенсивності (зазвичай 0 і 1). Це дає можливість суттєво зменшити кількість інформації, яка підлягає обробці. Результат цього скорочення може бути як достатньо успішним, так і невдалим. Невдалим результатом процедури бінаризації тут треба трактувати як появу різноманітних спотворень, наприклад, розриви границь, втрата деталізації, зашумленість та різноманітні спотворення (зокрема які виникають внаслідок неоднорідності фону тощо). Усі ці негативні результати бінаризації значно впливають на подальшу обробку зображення. Особливо актуальними вони є у випадку розв’язання задачі розпізнавання.

Існуючі сьогодні методи бінаризації можна поділити на дві групи:

- Порогові методи (нижня порогова обробка, верхня порогова обробка, порогова обробка з подвійними обмеженнями, неповна порогова обробка)
- Адаптивні методи (багаторівнева порогова обробка, локальна порогова обробка (метод Отсу, градієнтний, ентропійний та градієнтний методи), глобальна порогова обробка (методи Бернсена, Ейквіла, Ніблека, Яновиця і Брукштейна та ін.)

Принциповою відмінністю методів цих двох груп є існування єдиного порогового значення для усього зображення. Для методів адаптивної групи таке значення визначають у межах окремих ділянок зображень, що підвищує їх ефективність у випадку зображень, які характеризуються зміною освітленості. Очевидною, що у більшості випадків адаптивні методи є більш ресурсовитратними.

Основною перевагою бінаризації, особливо це стосується випадку порогових методів, є швидкість роботи її процедур. Проте, проблема вдалого вибору значення порогового критерію може нівелювати усі переваги процедури бінаризації. Так, наприклад, у багатьох практичних випадках не існує процедури автоматичного чи автоматизованого вибору бінаризаційного методу, а тому це доводиться здійснювати вручну.

Постановка задачі

Як вже зазначалось, без наявності фону задача сегментації є достатньо простою – її розв’язання реалізовується застосуванням адаптивного фільтрування за методом Отсу [32]. Проте наявність неоднорідного фону не дає можливості чітко відокремити бажану область від складного фону шляхом використання одного фільтрування. Це пояснюється тим, що при такому фільтруванні в

області сегментування обов'язково будуть захоплюватись різні ділянки фону. У роботі запропоновано підхід, який можна назвати двоетапним сегментуванням і який дає змогу доволі ефективно і точно виділяти шукані ділянки зображення для випадку неоднорідного фону.

Характеристика методу Отсу

Метод Отсу вважається одним із найефективніших методів бінаризації. Ефективність тут визначають за співвідношенням швидкості роботи процедури до якості результатів бінаризації. Результати бінаризації наведено на рис. 1 [33].



Вхідне зображення

Бінаризоване зображення за методом Отсу

Рис. 1. Результат бінаризації за методом Отсу

Метод Отсу належить до класу гістограмних методів, оскільки він ґрунтується на використанні гістограми розподілу значень функції інтенсивності. Сама гістограма складається із L проміжків і будується за імовірностями:

$$p_i = \frac{n_i}{N},$$

де N – загальна кількість пікселів зображення P ; n_i – кількість пікселів із значенням функції інтенсивності i .

За методом Отсу поріг T визначається із ідеї мінімізації дисперсії всередині піксельного класу, яка рівносильна максимізації дисперсії ($\sigma_m^2(k)$) між цими класами. Тому визначення порогового значення T можна записати так:

$$T = \arg \max_k \left(\sigma^2(k) \right) = \arg \max_k \left(w_1(k) w_2(k) [\mu_1(k) - \mu_2(k)]^2 \right),$$

де $k \in [0, L]$ – ціле число, яке розділяє діапазон значень функції інтенсивності на два класи, $w_0(k)$, $w_1(k)$, $\mu_0(k)$, $\mu_1(k)$ – імовірності (відносні частоти) та середні арифметичні класів, розділених порогом k , які визначаються так:

$$w_0(k) = \sum_{i=1}^k p_i; \quad w_1(k) = 1 - w_0(k)$$

$$\mu_0(k) = \sum_{i=1}^k \frac{i p_i}{w_0}; \quad \mu_1(k) = \sum_{i=k+1}^L \frac{i p_i}{w_1}.$$

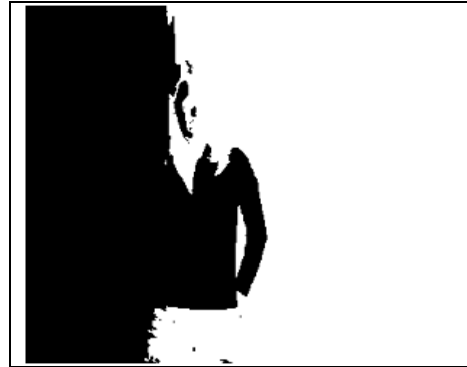
Основний недолік методу Отсу наведено на рис. 2 [32]. Цей недолік полягає у втраті деталізації зображення для випадку неоднорідного фону. Це добре проілюстровано на рис. 2, де вхідне зображення характеризується сильною зміною освітлення.

Властивості кольорових просторів

Розроблений метод ґрунтується на використанні різних кольорових просторів. Використання різних просторів уможливило фільтрування зображення за різними властивостями точок зображення. У випадках, якщо така фільтрація є можливою, то використання різних кольорних просторів суттєво спрощує самі процедури фільтрації.



Вхідне зображення



Бінаризоване за методом Отсу зображення

Рис. 2. Результат бінаризації за методом Отсу

Стандартна модель RGB дозволяє фільтрування лише за інтенсивностями червоної, синьої та зеленої колірних складових. Для відокремлення загальної яскравості піксела від кольору необхідні інші моделі. У цьому методі застосовують такі два колірні простори: HSV та YCbCr.

Колірний простір HSV описує колір піксела в базисі трьох величин: H – колірності, S – насиченості та V – яскравості. Він корисний тим, що дає змогу аналізувати яскравість точки без огляду на конкретний колір. Це полегшує детектування яскравих/неяскравих об'єктів, з насиченими/ненасиченими кольорами, без огляду на самі кольори.

Колірний простір YCbCr ґрунтується на такому базисі: Y – яскравість, подібна за визначенням до компоненти V моделі HSV); Cb – синій різницевий показник колірності; Cr – червоний різницевий показник колірності. На відміну від моделі HSV у моделі YCbCr не одна з трьох, а дві з трьох компонент відповідають за передавання колірності точки.

Послідовні використання різних кольорових просторів дають змогу сегментувати зображення за різними критеріальними ознаками.

Метод двоетапного сегментування зображення на основі методу Отсу

Метод двоетапного сегментування опишемо на такому практичному прикладі. Нехай існує деяке зображення, значну частину якого займає відносно світліший та однорідніший предмет. Наприклад, аркуш паперу. Крім нього, площину зображення можуть займати ще дві групи об'єктів: група зовнішніх об'єктів (тобто тих, які оточують центральний елемент), та група внутрішніх об'єктів (тобто тих, які знаходяться в межах(області) центрального об'єкта.

Тоді двоетапна сегментація зображення складатиметься із таких кроків:

Крок 1. Виділення центрального об'єкта (відкидання всього за межами центрального об'єкта);

При визначенні центрального об'єкта спершу зображення перетворюється з простору RGB на простір HSV, оскільки цей простір дозволяє відокремити кольорові характеристики пікселів від величини освітленості. Вважають, що центральний об'єкт має тенденцію зростання до вищої яскравості та спадання до нижчої насиченості. Він може бути і білого кольору (це не обов'язкове припущення). Низька насиченість кольору визначає урівноваженість усіх колірних каналів із вхідним каналом RGB. Тому для визначення центрального об'єкта проводимо фільтрування за каналами S та V . При цьому виділяють ті точки, для яких S менше ніж поріг, а V – більше.

Тобто шанси залишитись після фільтрування мають більше ті точки, які при зрівноважених каналах мають більші значення (екстремальний варіант – білий колір). Для кожного каналу (S та V) поріг розраховують окремо за алгоритмом Отсу. Використання цього методу є кроком до автоматизації алгоритму, тобто його адаптивності. Визначене за алгоритмом Отсу порогове значення ділить гистограму сірої палітри (а кожен кольоровий канал окремо можна так представити) на дві максимально рівнозначні частини, в кожній з яких варіація значень буде мінімальною. Тому цей метод якнайкраще підходить для випадків сегментації зображень на два класи (передній план і фон).

Отже, результатом фільтрації за методом Отсу буде деяка бінарна маска M_1 , над якою далі виконується морфологічна обробка, зокрема операція закриття отворів. Останню операцію проводять

для відокремленого видалення зовнішніх пікселів. У результаті залишиться зображення центрального об'єкта з об'єктами, які він містить.

Наклавши маску на первісне зображення у просторі RGB, отримаємо результат першого кроку.

Крок 2. Виділення зони інтересу (відкидання центрального елемента).

Завданням другого кроку є видалення зображення центрального елемента. При цьому операція видалення не повинна торкнутись об'єктів, які знаходяться в області центрального елемента. Для цього виконаємо перехід до простору YCbCr. Після цього відбувається фільтрування по каналу Y: відкидаються точки, більші за порогове значення. Саме порогове значення розраховується у спосіб, подібний до того, що використовувався на попередньому кроці. Таким чином відкидають світлішу частину зображення. Результатом фільтрування є маска M_2 .

Кінцеву маску M отримуємо накладенням отриманих раніше масок

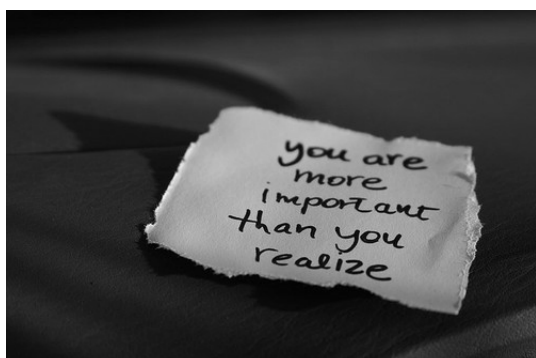
$$M = M_1 \cap M_2 .$$

Обидві маски є майже протилежні за своєю спрямованістю: перша відкидає світлі ділянки, а друга – темні. Однак отриманий результат зберігається завдяки морфологічній обробці маски M_1 . Результуюче зображення є результатом накладення маски M на початкове зображення.

Висновки

Результати використання запропонованого методу наведено на рис. 3 та 4. Оскільки метод розробляли в контексті загальної задачі розпізнавання тексту, то наведені результати стосуються зображень із текстом, зокрема рукописним.

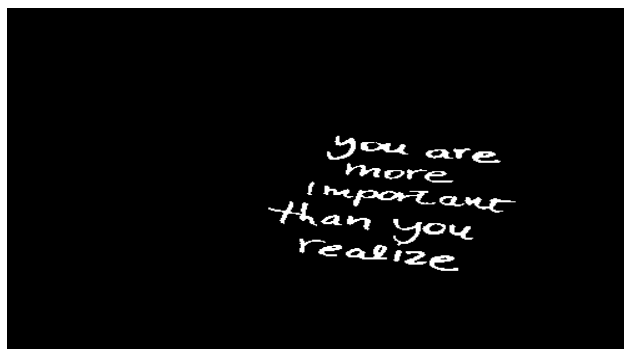
На рис. 3 результат роботи процедури, яка реалізовує алгоритм, побудований за розробленим методом, наведено поетапно. На рис. 4 наведено результати для випадку складніших вхідних зображень – мається на увазі складний (неоднорідний фон) і зашумлення центрального елемента. Незважаючи на це, результати роботи можна вважати задовільними.



Вхідне зображення

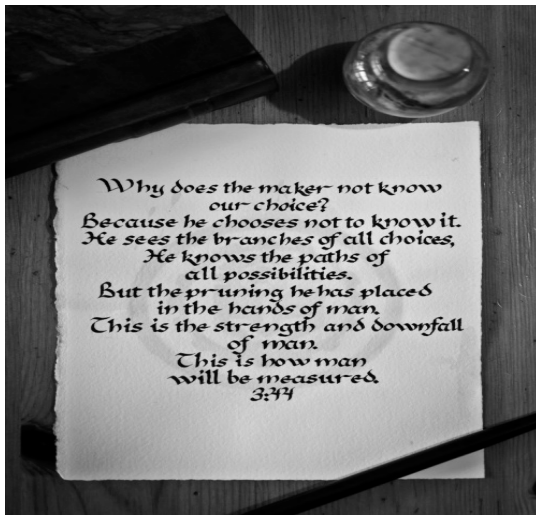


Бінаризоване зображення після першого кроку

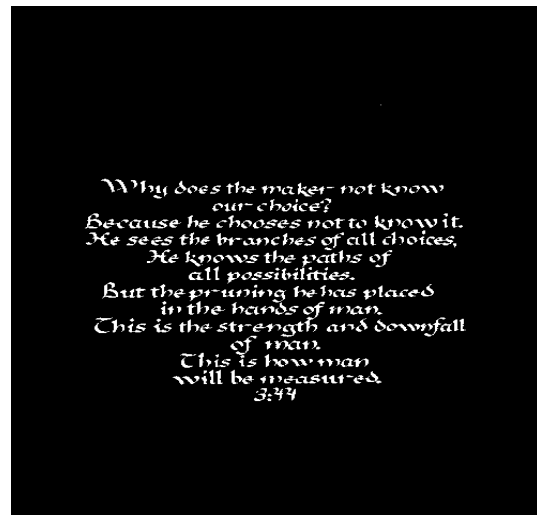


Бінаризоване зображення після другого кроку

Рис. 3. Приклад покрокового використання двоетапної сегментації для виділення тексту

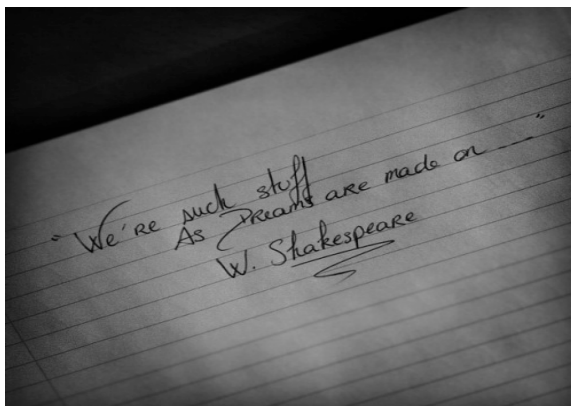


Вхідне зображення

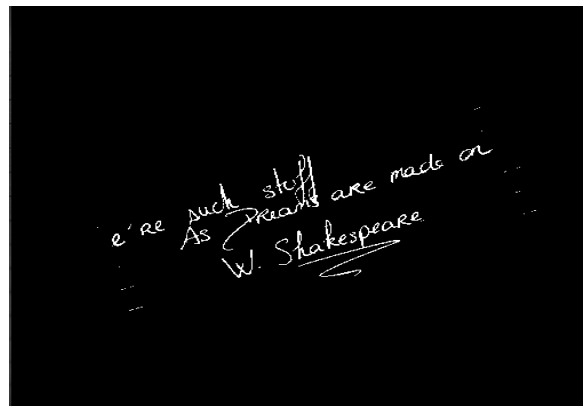


Сегментоване зображення

а) фон містить контрастні об'єкти, центральний об'єкт – прихований елемент



Вхідне зображення



Сегментоване зображення

б) фон є контрастним, центральний об'єкт є гармонійно зашумлений

Рис. 4. Приклад використання двоетапної сегментації для виділення тексту із складним фоном

На основі аналізу експериментальних результатів можна констатувати, що:

- метод є достатньо успішним навіть у випадках складного фону;
- метод є достатньо стійким до зашумлень центрального елемента. У випадку гармонійного періодичного чи майже періодичного зашумлення центрального елемента (рис. 4, б), можуть виникати незначні артефакти.

Основними перевагами розробленого методу, безумовно, можна вважати високу швидкість і низькі вимоги до обчислювальних ресурсів. Це визначається насамперед використанням достатньо ресурсоневиногливих алгоритмів конвертації в колірні простори і алгоритму Отсу.

1. Mehmet Sezgin, Bulent Sankur Survey over image thresholding techniques and quantitative performance evaluation *Journal of Electronic Imaging* 13(1), 146–165 (January 2004). 2. Sankur B., Sezgin M. A Survey Over Image Thresholding Techniques And Quantitative Performance Evaluation, (accepted), *Journal of Electronic Imaging*, 13(1), 146- 165, January, 2004. 3. H. D. Cheng, Y. H. Chen, and Y. Sun A novel fuzzy entropy approach to image enhancement and thresholding, *Signal Process.* 75, 277–301, 1999. 4. J. Z. Liu and W. Q. Li, The automatic thresholding of gray-level pictures via two-dimensional Otsu method, *Acta Automatica Sin.* 19, 101–105, 1993. 5. T. Pun, A new method for gray-level picture threshold using the entropy of the histogram, *Signal Process.* 2-3!, 223–237, 1980. 6. T. Pun

Entropic thresholding: A new approach, *Comput. Graph. Image Process.* 16, 210–239. – 1981.
 7. J. N. Kapur, P. K. Sahoo, and A. K. C. Wong A new method for gray-level picture thresholding using the entropy of the histogram, *Graph. Models Image Process.* 29, 273–285, 1985. 8. C. Chang, K. Chen, J. Wang, and M. L. G. Althouse, A relative entropy based approach in image thresholding, *Pattern Recogn.* 27, 1275–1289, 1994. 9. N. Otsu A threshold selection method from gray level histograms *IEEE Trans. Syst. Man Cybern. SMC-9*, 62–66, 1979. 10. D. E. Lloyd Automatic target classification using moment invariant of image shapes, *Technical Report, RAE IDN AW126, Farnborough, UK, Dec. 1985.*
 11. J. Kittler and J. Illingworth Minimum error thresholding *Pattern Recogn.* 19, 41–47, 1986.
 12. J. Kittler and J. Illingworth On threshold selection using clustering criteria *IEEE Trans. Syst. Man Cybern. SMC-15*, 652–655, 1985. 13. C. V. Jawahar, P. K. Biswas, and A. K. Ray Investigations on fuzzy thresholding based on fuzzy clustering *Pattern Recogn.* 10, 1605–1613, 1997. 14. C. M. Tsai and H. H. Lee Binarization of color document images via luminance and saturation color features *IEEE Trans. Image Process.* IP-11, 434–451, Apr. 2002. 15. L. Hertz and R. W. Schafer Multilevel thresholding using edge matching *Comput. Vis. Graph. Image Process.* 44, 279–295, 1988. 16. W. H. Tsai Moment-preserving thresholding: A new approach *Graph. Models Image Process.* 19, 377–393, 1985. 17. S. C. Cheng and W. H. Tsai A neural network approach of the moment-preserving technique and its application to thresholding *IEEE Trans. Comput. C-42*, 501–507, 1993. 18. L. O’Gorman Binarization and multithresholding of document images using connectivity,” *Graph. Models Image Process.* 56, 494–506, 1994. 19. R. L. Kirby and A. Rosenfeld A note on the use of ~gray level, local average gray level! space as an aid in threshold selection *IEEE Trans. Syst. Man Cybern. SMC-9*, 860–864, 1979. 20. A. Rosenfeld and R. Smith Thresholding using relaxation *IEEE Trans. Pattern Anal. Mach. Intell. PAMI-3*, 598–606, 1981. 21. A. Y. Wu, T. H. Hong, and A. Rosenfeld Threshold selection using Survey over image thresholding techniques... *Journal of Electronic Imaging / January 2004 / Vol. 13(1) / 167 quadrees* *IEEE Trans. Pattern Anal. Mach. Intell. PAMI-4*, 90–94, 1982. 22. N. Ahuja and A. Rosenfeld A note on the use of second-order graylevel statistics for threshold selection *IEEE Trans. Syst. Man Cybern. SMC-5*, 383–388, 1975. 23. J. Sauvola and M. Pietaksinen Adaptive document image binarization *Pattern Recogn.* 33, 225–236, 2000. 24. J. M. White and G. D. Rohrer Image thresholding for optical character recognition and other applications requiring character image extraction *IBM J. Res. Dev.* 27, 400–411, 1983. 25. Y. Yasuda, M. Dubois, and T. S. Huang Data compression for check processing machines *Proc. IEEE* 68, 874–885, 1980. 26. S. D. Yanowitz and A. M. Bruckstein A new method for image segmentation *Comput. Graph. Image Process.* 46, 82–95, 1989. 27. D. Shen and H. H. S. Ip A Hopfield neural network for adaptive image segmentation: An active surface paradigm *Pattern Recogn. Lett.* 18, 37–48, 1997. 28. S. K. Pal and A. Rosenfeld Image enhancement and thresholding by optimization of fuzzy compactness *Pattern Recogn. Lett.* 7, 77–86, 1988. 29. A. S. Abutaleb Automatic thresholding of gray-level pictures using two-dimensional entropy *Comput. Vis. Graph. Image Process.* 47, 22–32, 1989. 30. F. Deravi and S. K. Pal Grey level thresholding using second-order statistics *Pattern Recogn. Lett.* 1, 417–422, 1983. 31. Бинаризация изображений: алгоритм Брэдли(<https://habrahabr.ru/post/278435/>). 32. Сегментация изображения (<https://habrahabr.ru/post/128768/>).33.[https://www.wikiwand.com/ ru/%D0 %9C% D0 % B5 %D1 %82 %D0 %BE%D0 %B4_ %D0 %9E%D1 %86 %D1 %83](https://www.wikiwand.com/ru/%D0%9C%D0%B5%D1%82%D0%BE%D0%B4_%D0%9E%D1%86%D1%83).