

УДК 004.932

К.т.н., доцент Сулема Є.С., студент Гуренко О.А.

Національний технічний університет України  
«Київський політехнічний інститут»

## МОДИФІКОВАНИЙ АЛГОРИТМ ВИЯВЛЕННЯ ТЕКСТОВОЇ ІНФОРМАЦІЇ У ЗОБРАЖЕННІ

### Abstract

*Yevgeniya S. Sulema, assoc. prof., PhD; Gurenko Oleg, student  
Modified algorithm for text information detection in an image*

*This paper concerns the task of text detection on image. Existing algorithms of text detection are studied and discussed. The modified algorithm based on pixels contrast difference with the analysis of potential text areas characteristics is proposed. The algorithm is tested on a large number of images. The strengths and weaknesses of the algorithm are identified and discussed in the article.*

### Вступ

Задача виявлення тексту, що міститься у зображенні, є однією з задач розпізнавання образів. Вона стосується багатьох сфер людської діяльності: від військової справи й систем автоматичної ідентифікації до кіноіндустрії та телебачення. На сьогодні запропоновано багато способів виділення тексту із зображення [1-5], більшість з яких ґрунтується або на аналізі статистичних характеристик фрагментів зображення, або на попіксельній обробці зображення.

При виконанні аналізу статистичних характеристик до уваги береться зміна цих характеристик при переході від одного фрагмента до наступного при їх відрядковому скануванні. При цьому визначаються ознаки, що ґрунтуються на статистиках першого порядку: гістограма першого порядку, асиметрія, ексцес, ентропія тощо.

При виконанні попіксельної обробки зображення для кожного пікселя виконується аналіз сусідніх пікселів у певному радіусі. Для локалізації тексту використовують фільтрацію, метою якої є виділення границь об'єктів (*Edge Detection*), виділення кутів (*Corner Detection*) і виділення текстур тексту (*Texture Detection*).

Обидва підходи можуть бути реалізовані багатьма способами, причому наразі жоден з них не дозволяє досягти 100% точності, отже

задача пошуку нових способів виявлення тексту в зображенні залишається актуальною.

## Постановка задачі

Задача, вирішення якої розглядається в рамках даної статті, полягає у попіксельному аналізі зображення з метою визначення його фрагментів, що містять текст (літери та цифри). Передбачається, що отриманий результат буде використаний для захисту знайденого тексту від несанкціонованого доступу при передаванні зображення, що містить текст, по відкритих каналах зв'язку.

## Опис алгоритму

Запропонований алгоритм виявлення текстової інформації на зображенні ґрунтується на тому, що звичайно колір тексту є контрастним до кольору тла. Тому пропонується визначати та аналізувати різницю інтенсивності кольору сусідніх пікселів у зображенні. Загальна схема процедури виявлення тексту у зображенні представлена на рис. 1.

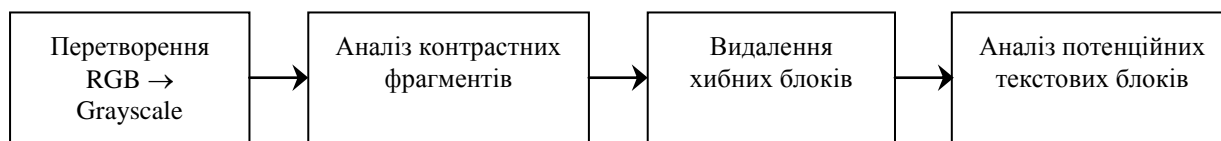


Рис.1. Схема процедури виявлення тексту у зображенні

Алгоритм, що пропонується, передбачає наступні дії.

1. Перетворення кольорового зображення у відтінки сірого кольору. Зображення як структура даних являє собою тривимірний масив  $A$  розміром  $N \times M \times 3$ . Після перетворення отримуємо двовимірний масив  $A'$  розміром  $N \times M$ . Перетворення виконується за формулою:

$$A'_{ij} = A_{ij1} \cdot k_r + A_{ij2} \cdot k_g + A_{ij3} \cdot k_b,$$

де  $k_r, k_g, k_b$  – коефіцієнти перетворення ( $k_r = 0.299, k_g = 0.587, k_b = 0.114$ ),  $i, j$  – позиція поточного пікселя у зображенні ( $i=1 \dots N, j=1 \dots M$ ).

Дане перетворення потрібно для посилення контрастності зображення, що сприятиме точності знаходження тексту.

2. Попіксельний обхід зображення та аналіз контрастності в околі поточного пікселя:
  - якщо різниця значень інтенсивності відтінку сірого кольору перевищує заданий поріг  $t$ , то сусідній піксель вважається

контрастним до даного; в результаті отримуємо матрицю  $I$  контрастних пікселів:

$$I_{ij} = q,$$

де  $q$  – кількість контрастних пікселів, сусідніх до даного ( $0 \leq q \leq 8$ );

3. Перетворення матриці контрастних пікселів у булеву матрицю контрастних пікселів:

$$B_{ij} = \begin{cases} 1, & 3 \leq I_{ij} \leq 8 \\ 0, & I_{ij} < 3 \end{cases}.$$

4. Аналіз булевої матриці контрастних пікселів  $B$  з метою виявлення довгих послідовностей одиниць, що стоять підряд або чергуються з невеликою кількістю нулів, та формування матриці потенційних текстових блоків  $T$ . Аналіз матриці відбувається по рядках векторами довжини  $l_1$ . Емпіричним шляхом було визначено, що для виявлення декількох літер тексту одночасно потрібно, щоб виконувалась умова  $35 \leq l_1 \leq 45$ . Далі визначаємо кількість одиниць у векторі довжини  $l_1$ :

- якщо знайдено більше, ніж  $l_2$ , одиниць, що стоять підряд, то скоріш за все вони відповідають лінії у зображенні, оскільки літери в словах звичайно утворюють переривчасту лінію; такі довгі послідовності одиниць ігноруються;
- якщо за один крок у матриці  $B$  знайдено більше, ніж  $l_3$ , одиниць, то можливо виявлено текст, тому цьому фрагменту матриці  $B$  ставиться у відповідність еквівалентний за розміром фрагмент з одиниць у матриці  $T$ .

В ході експериментів було визначено, що найбільшу точність виявлення елементів тексту можна досягти при  $l_2 = 25$  та  $l_3 = 16$ .

5. Для окреслення прямокутної області, що потенційно містить текст, використовується хвильовий алгоритм. Він передбачає повний обхід матриці потенційних текстових блоків  $T$ . В результаті отримуємо дві координати, за якими будується прямокутна область. Процедура повторюється для кожного виділеного блоку.

Наприкінці виконання алгоритму отримуємо прямокутні області, що «накривають» текстові блоки. Даний алгоритм забезпечує виділення тексту, розташованого горизонтально або під кутом до 7 градусів, а також тексту, що розташований на уявній хвилястій лінії. Для відкидання областей, що насправді не містять тексту, отримані прямокутні області аналізуються щодо їх відповідності особливим характеристикам тексту. Наприклад, текстовим блоком вважатиметься той, що має висоту від 6 до

44 пікселів та ширину вдвічі більшу, за висоту; площа потенційного текстового об'єкта більша, ніж половина площини всього блока.

Результатом виконання алгоритму є зображення, на якому виділено прямокутні області, що з високою ймовірністю містять текстову інформацію.

### Аналіз запропонованого алгоритму

Аналіз алгоритму виконувався на зображеннях з різним тлом й текстом різного кольору, величини та орієнтації. Результати аналізу наведені у табл. 1.

Таким чином, алгоритм характеризується ймовірністю виявлення тексту в зображенні, близькою до 0,9.

Алгоритм має обмеження щодо виявлення тексту з наступними характеристиками:

- текст під кутом понад 7 градусів;
- короткі слова, менші за 4 символи;
- розмір букв, більший за 44 пікселів;
- розмір букв, менший за 6 пікселів;
- колір букв, близький до кольору тла;
- тло з великою кількістю дрібних деталей.

Час обробки зображення варіюється у межах 2-4 секунди для зображень розміром до 1024x768 пікселів.

Таблиця 1

Показники роботи алгоритму для різних типів зображень та тексту

| Тип зображення                | Кількість проаналізованих зображень | Кількість текстових елементів на зображеннях | Кількість знайдених елементів | Кількість хибних захоплень | Відсоток виявлення |
|-------------------------------|-------------------------------------|--|-------------------------------|----------------------------|--------------------|
| Біле тло та звичайний текст   | 20                                  | 72   | 66                            | 0                          | 92%                |
| Біле тло та текст під нахилом | 20                                  | 44   | 40                            | 0                          | 91%                |
| Строкате тло                  | 20                                  | 60   | 51                            | 4                          | 85%                |
| Фотографії, що містять текст  | 20                                  | 46   | 40                            | 11                         | 87%                |

Шукані характеристики тексту можна змінювати під певні класи задач, тим самим збільшуючи ймовірність правильного виявлення текстових блоків.

## Висновки

Запропонований алгоритм ґрунтується на попiксельнiй обробцi зображення. Вiн дозволяє визначати областi зображення, що мiстять текстову iнформацiю. Ймовiрнiсть правильного знаходження тексту складає близько 0,9. Якiсть результату роботи алгоритму залежить вiд вигляду тла та характеристик шрифту, яким надруковано текст, а також iх кольорiв: чим вища контрастнiсть тексту вiдносно тла, тим кращий результат роботи алгоритму.

Час роботи алгоритму складає близько 3-х секунд для зображень великого розмiру (високої роздiльностi). Даний алгоритм має обмеження щодо вигляду тла та тексту, проте його роботу можна оптимiзувати шляхом урахування особливостей класу оброблюваних зображень.

## Лiтература

1. *Горелик А. Л.* Методы распознавания [Текст] / А. Л. Горелик, В. А. Скрипкин – М. : Высшая школа, 1989. – 208 с.
2. Математические методы распознавания образов [Текст] // Сборник докладов 10-й Всероссийской конференции. – М. : АЛЕВ-В, 2001. – 342 с.
3. *Загоруйко Н.Г.* Методы распознавания и их применение [Текст] / Н. Г. Загоруйко. – М. : Советское радио, 1972. – 208 с.
4. Anthimopoulos, M. A two-stage scheme for text detection in video images [Text] / M. Anthimopoulos, B. Gatos, I. Pratikakis // Image and Vision Computing. – Vol. 28, Issue 9. – P. 1413-1426.
5. *Liu Q.* Stroke filter for text localization in video images [Text] / Q. Liu, C. Jung, S. Kim, Y. Moon, J. Kim // Proc. IEEE Conf. Image Processing. – P. 1473-1476.