

УДК 004.932.2

О.В. Гороховатський

Харківський національний економічний університет імені Семена Кузнеця, Харків

ОСОБЛИВОСТІ РОЗПІЗНАВАННЯ ЗОБРАЖЕНЬ СИМВОЛІВ ІЗ ВИКОРИСТАННЯМ ЛІНІЙНИХ ОПИСІВ ТА КОРЕКЦІЇ РЕЗУЛЬТАТІВ

Розглянуто впровадження методу побудови лінійних описів для вирішення задачі розпізнавання тексту. Лінійний опис будується в процесі інтерпретації структурних зв'язків об'єкту. Досліджено проблемні аспекти використання запропонованого методу, розглянуто можливі методи корекції тексту після розпізнавання.

Ключові слова: символ, блок, лінійний опис, стан, відстань Левенштейна, корекція.

Вступ

Розпізнавання тексту залишається актуальною проблемою в галузях обробки зображень та комп'ютерного зору через масове впровадження автоматичного програмного забезпечення для вирішення подібних завдань. Не дивлячись на те, що на поточний момент існують досить потужні системи розпізнавання тексту, вони є досить залежними від початкових умов та пов'язані зі значними часовими та обчислювальними витратами при створенні ознак, придатних для розпізнавання, та їх аналізом та порівнянням.

Використання ж структурних та чисельних ознак, які характеризуються простотою побудови в реальних задачах, залишається недооціненим.

Статтю присвячено дослідженню впровадження методу побудови лінійного опису зображення символу для вирішення практичної задачі розпізнавання. Опис створюється за допомогою операції сканування значущих ліній зображення і їх подальшої спеціальної обробки, що загалом не потребує значних обчислювальних операцій.

Ідея методу полягає в реалізації розпізнавання об'єктів за формою та може бути легко адаптована до використання при роботі з будь-якими об'єктами довільної форми.

Метою даної статті є дослідження можливостей практичного використання лінійного опису, дослідження якості роботи запропонованого методу, дослідження можливих методів корекції результатів розпізнавання.

Попередня обробка зображення

Попередня обробка зображення дозволяє покращити його якість та зробити його аналіз та обробку більш зручними та стабільними.

Традиційним етапом обробки зображення текстових фрагментів є бінаризація, тобто зведення зображення лише до двох градацій яскравості, котрі відповідають пікселям тексту та пікселям фону.

Для прискорення процесу обробки запропоновано не виконувати попередню обробку всього зображення, а обробляти лише значущу його частину. Це досягнуто із використанням нижче наведеної послідовності кроків.

1. Розбиття вихідного цілісного зображення на сукупність зображень окремих літер за допомогою послідовного горизонтального та вертикального проектування та аналізу проміжків та мінімальних значень одновірної функції-проекції.

2. Розмиття зображення за допомогою матричного фільтру розміру 3×3 , що складається з одиниць.

3. Медіанна фільтрація зображення маскою розміру 3×3 .

4. Бінаризація зображення із використанням адаптивного порогового значення Отсу [1].

Лінійні описи зображення

Розглянемо зображення $I(x, y)$ як сукупність горизонтальних ліній $L = \{l_1, l_2, \dots, l_n\}$ із кроком Δl [2 – 5].

Під блоком b матимемо на увазі послідовність пікселів чорного кольору, чергування котрих утворює лінію $l = \bigcup_{i=1}^m b_i$. Значення m визначається практично і відповідає максимально можливій кількості блоків, наприклад, для латинського алфавіту можна прийняти $m = 4$.

Нехай $S^i = \{s_1, s_2, \dots, s_{k_i}\}$, $i = \overline{1, m}$ – множина станів, котрі може бути утворено комбінаціями блоків в кожній лінії, k_i – кількість можливих станів у множині S^i . Наприклад, наявність єдиного блоку в лінії може вказувати на один із станів множини

$S^1 = \{\text{"vertical line"}, \text{"horizontal line"}, \text{"diagonal line"}\}$.

Таким чином, кожна лінія l вихідного зображення відображається в один з можливих станів:

$L \Rightarrow s$, відповідно, множина всіх ліній зображення формує набір станів

$$L \Rightarrow C = \bigcup s,$$

котрий і є лінійним описом зображення [2 – 5].

Разом із лінійним описом введено визначення профілю зображення, тобто відображення

$$L \Rightarrow P = \bigcup S.$$

Такі профілі не задовольняють критерію однозначної відповідності одному з еталонних класів і можуть бути використані лише в якості додаткового елемента розпізнавання.

Особливості порівняння описів

Розпізнавання доцільно будувати на двох лінійних описах зображення – горизонтальному та вертикальному. Міра подібності лінійних описів будується із використанням відстані Левенштейна [6] та вагових коефіцієнтів нормалізації:

$$d(C_1, C_2) = w_C d_{Lev}(C_1, C_2) + w_P d_{Lev}(P_1, P_2), \quad (1)$$

де C_1, C_2 – порівнювальні описи, P_1, P_2 – профілі цих описів, $w_C = 1$, $w_P = 0.3$ – вагові коефіцієнти відстаней між описами та профілями відповідно.

Результати експериментів

Експерименти з розпізнавання проводилися на навчальній множині, елементи котрої містили 407 екземпляри прописних та строкових літер латинського алфавіту та 832 екземпляри спеціальних символів та цифр шрифту Arial кегль 14 пунктів, формуючи таким чином 50 еталонних класів. Навчальний текст біло надруковано та скановано з роздільною здатністю 300 dpi.

Для оцінювання якості розпізнавання було використано тестовий фрагмент тексту розміром близько 2500 літер включно з пропусками, роздрукованому та сканованому з роздільною здатністю також 300 dpi.

Розпізнавання із використанням відстані (1) на базі профілів та лінійних описів зображення дозволяє досягти показника якості 76% із помилками у близько 100 слів з 426 (рис. 1).

До типових помилок слід віднести наступні:

– помилки через некоректне виділення літер (рис. 1, перший рядок);

– помилки через близькість профілів або лінійних станів. Прикладами таких помилок є: «rface» замість «place», «sfars» замість «stars», «!sn,t» замість «isn't».

Використання канонічної відстані Левенштейна або варіації (1) для пошуку найбільш вірогідного коректного слова замість результату розпізнавання з помилкою не є достатньо ефективним.

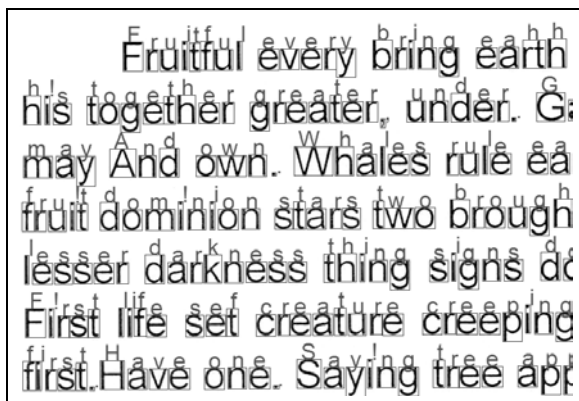


Рис. 1. Результат розпізнавання без корекції

Наприклад, для помилкового слова «rface» відстань Левенштейна є однаковою до слів «rface» та «rase».

Тож в основу методів корекції покладено апріорну інформацію про можливі помилки.

Так, коефіцієнти для обрахування відстані Левенштейна змінено для операцій видалення, заміни та додавання символу.

Виправлення однієї з відомих типових помилок, наприклад, «rface» у «rface» відбувається зі зменшеним коефіцієнтом, ніж виправлення менш типових помилок.

Така стратегія дозволяє налаштувати корекційні модулі відповідно до особливостей запропонованого алгоритму.

Найбільш проблемним є розпізнавання наступних комбінацій літер: «l» – «d», «>» – «>», «'» – «>», «!» – «i», «t» – «f», «g» – «9».

Розглянемо методи корекції результатів розпізнавання. Введемо первинну корекцію, яка буде полягати в тому, що слова, які містять спеціальні символи, будуть перевірені та виправлені із використанням словника.

Первинна корекція дозволяє досягти показника близько 88% (50 слів з 426 розпізнано з помилками, рис. 2). Порівняно із рис. 1 можна побачити виправлення в словах «First», «fruit», «dominion».

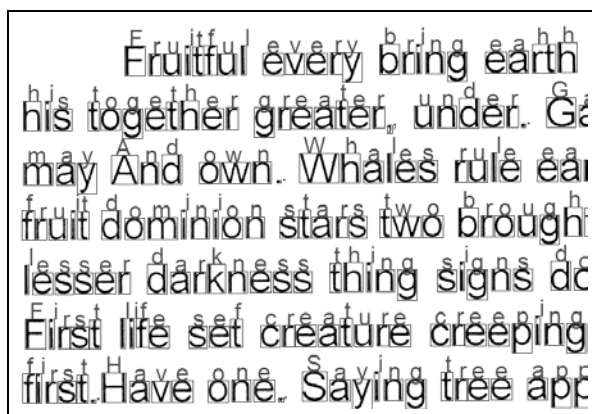


Рис. 2. Результат розпізнавання з первинною корекцією

Розглянемо впровадження повноцінної пост-корекції результатів розпізнавання, суть якої полягає в тому, що словниковий пошук виконується для кожного зі слів, що входять в результат розпізнавання.

Повноцінна пост-корекція дозволила досягти показника 93% (21 слово з 426 розпізнано з помилками, рис. 3).

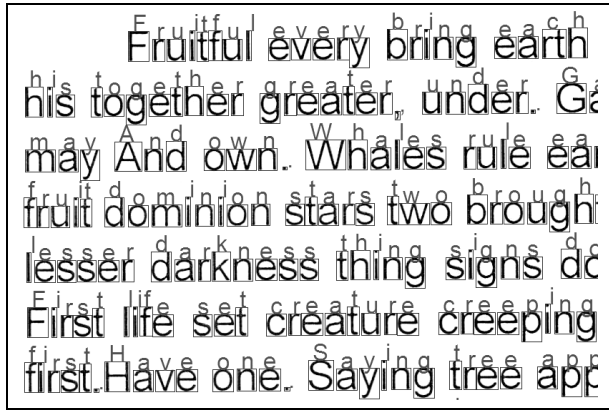


Рис. 3. Результат розпізнавання з повною корекцією

Порівнявши цей результат з зображенням на рис. 2, можна побачити виправлення в словах «set» та «each».

Висновки

Дослідження присвячено впровадженню для вирішення практичної задачі розпізнавання текстової інформації за його лінійним описом.

Відображення зображення в лінійний опис досягнуто шляхом структурного аналізу символу зображення за допомогою сканування лініями.

Науковою новизною роботи є методика використання лінійного опису та профілю зображення для вирішення задачі розпізнавання, а також процедури корекції результатів розпізнавання,

адаптовані для нівелювання відповідних недоліків методу.

В якості подальших досліджень можна розглядати шляхи прискорення як відповідно розпізнавання, так і методів корекції результатів, наприклад, за рахунок використання більш сучасних технічних методів організації та обробки даних, а також покращенням відповідних методів.

Список літератури

1. Otsu N. A threshold selection method from gray-level histograms / N. Otsu // *IEEE Transactions on Systems, Man and Cybernetics*. – 1979. – Vol. 9. – P. 62-66.
2. Гороховатский А.В. Распознавание изображений символов по их линейному описанию / А.В. Гороховатский, Е.О. Передрий // *Интеллектуальні системи прийняття рішень і проблеми обчислювального інтелекту (ISDMCI'2012): Зб. наук. праць міжн. наук. конф.: Євпаторія, 17-22 травня 2012.* – Херсон: ХНТУ. – С. 340-342.
3. Гороховатский А.В. Распознавание изображений символов на основе линейного описания структурных характеристик / А.В. Гороховатский, Е.О. Передрий // *Системы обработки информации*. – X.: ХУ ПС, 2012. – Вып. 7(105). – С. 203-206.
4. Gorokhovatskyi O. Primary Symbol Recognition Using Scanline Structural Descriptions / O. Gorokhovatskyi, O. Peredrii // *6th International Academic Conference of Young Scientists "Computer Science and Engineering 2013" (CSE-2013): Lviv, November 21-23, 2013.* – P. 156-157.
5. Гороховатский А. Методы улучшения производительности процедуры сравнения признаков при распознавании изображений символов / А.В. Гороховатский // *Системы обработки информации*. – X.: ХУ ПС, 2015. – Вып. 1(126). – С. 20-22.
6. Левенштейн В.И. Двоичные коды с исправлением выпадений, вставок и замещений символов / В.И. Левенштейн // *Доклады Академии Наук СССР*. – 1965. – № 163.4. – С. 845-848.

Надійшла до редколегії 1.02.2016

Рецензент: д-р техн. наук, проф. Є.П. Путятін, Харківський національний університет радіоелектроніки, Харків.

ОСОБЕННОСТИ РАСПОЗНАВАНИЯ ИЗОБРАЖЕНИЙ СИМВОЛОВ С ИСПОЛЬЗОВАНИЕМ ЛИНЕЙНЫХ ОПИСАНИЙ И КОРРЕКЦИИ РЕЗУЛЬТАТОВ

А.В. Гороховатский

Рассмотрено внедрение метода построения линейных описаний для решения задачи распознавания текста. Линейное описание строится в процессе интерпретации структурных связей объекта. Исследованы проблемные аспекты использования предложенного метода, рассмотрены возможные методы коррекции текста после распознавания.

Ключевые слова: символ, блок, линейное описание, состояние, расстояние Левенштейна, коррекция.

CHARACTERISTIC OF SYMBOL IMAGE RECOGNITION USING LINEAR DESCRIPTIONS AND CORRECTION OF RESULTS

O.V. Gorokhovatskyi

Implementation of the a method of construction of linear descriptions for solving the problem of text recognition is considered. Linear descriptions are built during an interpretation of structural connections of the object. Specific problematic aspects of the proposed method are investigated, possible methods of text post-recognition correction are presented.

Keywords: symbol, block, linear description, condition, Levenshtein distance, correction.